AFOSR - TR - 77 - 0640

COMPUTER SCIENCE
TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

TR-515
F44620-72C-0062

March 1977

TECHNIQUES FOR SEGMENTING INFRARED CLOUD COVER IMAGES

JoAnn Parikh
Azriel Rosenfeld
Computer Science Center
University of Maryland
College Park, MD 20742

ABSTRACT

This report investigates three techniques for segmenting cloud cover images into regions of homogeneous cloud type. Two of these techniques select thresholds based on an analysis of the edge strengths of the borders of the above-threshold connected components (or of the coldest such component). The third technique selects thresholds based on cluster analysis of the infrared histogram, combined with a statistical feature analysis of the clusters in the image domain.

The support of the Directorate of Mathematical and Information Sciences, U. S. Air Force Office of Scientific Research, under Contract F44620-72C-0062, is gratefully acknowledged, as is the help of Ms. Shelly Rowe.

1. Introduction

This report investigates three techniques for segmenting cloud cover images:

1) Segmentation by selection of thresholds based on analysis of the average border edge strength of the coldest connected component.

2) Segmentation by selection of thresholds based on analysis of average border edge strengths of all connected components.

3) Segmentation by selection of temperature interval clusters based on a cluster analysis of the infrared histogram combined with a statistical feature analysis of the clusters in the image domain.

A description of the data set for the segmentation study in this report is presented in Section 2. A comparison of thresholds suggested by analysis of average border edge strength of the coldest connected component with thresholds suggested by analysis of average border edge strength is the subject of Section 3. Average border edge strength was calculated by applying the Roberts gradient operator to the infrared image data. Average border edge strength features were calculated in Section 4 by applying three different edge operators. Threshold selection results were compared for all three edge operators. Section 5 approaches the segmentation problem in terms of clustering analysis rather than edge analysis. Features in the infrared image

domain were then calculated on the clusters to identify and/or merge the clusters into cloud type and/or cloud layer objects. Section 6 describes a test of the cluster analysis approach on a data base of 107 cloud cover windows.

2. Data Set

The data samples selected for the segmentation study presented in this report consist of the first twelve samples of Data Set III (described in [1]). The large size infrared windows of Data Set III (64x64 arrays of infrared picture points) were used for each of the segmentation techniques. Twelve visible windows (64x32 arrays of visible picture points) corresponding to the central portions of the twelve large size infrared windows were used for only one of the segmentation techniques -- cluster analysis based on a visible brightness feature (Section 5). Visible brightness features were calculated from each of the visible windows in an attempt to classify cluster cloud type. Analysis of cluster cloud type based on an infrared edge strength feature is also discussed in Section 5.

The twelve data samples consist of three samples of "cumulonimbus" cloud type, four samples of "low" cloud type, and five samples of "mix" cloud type. One of the "low" samples, Sample Number 5, in this sample set for the pilot segmentation study contains primarily middle clouds. One of the "mix" samples, Sample Number 4, contains dense cirrus produced by vigorous deep convection, whereas the other four "mix" samples, Sample Numbers 8-11, were selected from regions in which there were no neighboring "cumulonimbus" samples. The geographical location (see Figures 1 and 2 of [1]) and classification for each of the twelve samples are given in Table 1.

3. Connected Component Analysis

For each of the twelve samples, the 64x64 array of infrared temperature readings was contoured for every temperature threshold T_i occurring in the sample into sets of connected components $C_{i1},\ldots,C_{im_i}$. Each connected component $C_{i1},\ldots,C_{im_i}$ consisted only of points with infrared temperature readings colder than T_i. A set of contiguous infrared picture points C_{ir} was defined as a connected component for a given temperature threshold T_i if all points contained in C_{ir} had infrared temperature readings greater than T_i (i.e., colder) and for any two points p and q in C_{ir}, there existed a path of points in C_{ir} between p and q such that each point along the path was either horizontally, vertically, or diagonally adjacent to the preceding point along the path. A point of a connected component C_{ir} was said to be a border point of C_{ir} if at least one of its vertical or horizontal neighbors had an infrared temperature reading less than or equal to T_i. An example of the set of connected components corresponding to a temperature threshold of 132 for a 4x6 array of infrared picture points is shown below. Border points for each component are underlined.

<u>137</u>	<u>142</u>	<u>144</u>	<u>143</u>	
<u>138</u>	147	156	<u>165</u>	COMPONENT 1
132	<u>142</u>	<u>147</u>	<u>159</u>	
118	121	124	132	
117	119	129	<u>137</u>	COMPONENT 2
124	127	<u>136</u>	<u>146</u>	

A detailed discussion of the concepts of adjacency and connectedness can be found in Rosenfeld and Kak [2].

The purpose of the experiments in this section was to assess the value of isolating, labeling, and computing border edge strength for each of the spatially distinct connected components whose temperature readings fell within the same range. It was assumed that, if each connected component for a given threshold represented a distinct area of cloud-type data, the border edge strength for a particular connected component might be more indicative of an appropriate threshold value than the border edge strength averaged over all the connected components. The particular connected component which was of primary interest was the connected component containing the coldest temperature reading. In the case of "cumulonimbus" samples, this component would represent the cloud-type area (or object) in which the most vigorous convection was occurring. In the case of "mix" samples consisting of thin cirrus mixed with low clouds, this component would probably best represent the thin cirrus portion of the sample. Figures 1-12 illustrate the difference for each threshold value between border edge strength values for the component containing the coldest temperature (solid line graph) and the border edge strength averaged over all connected components (dotted graph).

The edge strength operator used to determine border edge strength was the Roberts gradient. For any 2x2 array

of points $\begin{matrix} A & B \\ C & D \end{matrix}$, the value of the edge strength at point A was defined as

$$\max(|A-D|, |C-B|).$$

The edge strength of a connected component C_{ir} was then calculated as the sum of the edge values for all border points of C_{ir} divided by the number of border points of C_{ir}. The border edge strength for all components corresponding to a given threshold T_i was calculated by finding the sum of edge values for all border points of the connected components associated with T_i and then dividing by the total number of border points.

An analysis of the graphs in Figures 1-3 for "cumulonimbus" samples reveals that it is very difficult to select a temperature threshold to contour the cumulonimbus portion of the sample either by examining the pattern of edge strength values for the connected component containing the coldest temperature or the pattern of edge strength values averaged over all connected components. For Figures 1-3, an attempt was made by looking at the graphs and the edge strength data from which the graphs were drawn to find the best temperature threshold for separating the cumulonimbus portion of the sample from the cirrus and low cloud portion by scanning the cold temperature portion (right-hand side of the graphs) for either high edge strength values or sharp increases in edge strength values from warmer to colder temperatures.

In Figure 1, there was only one connected component for infrared readings from 166 to 205. The edge strength values generally decreased throughout this range. No specific threshold could be found from the graph for separating the cumulonimbus portion of Sample Number 1 (shown in Figure 1) from the cirrus portion.

The situation was somewhat different for Figure 2. Although total edge strength values decreased in general over a range of temperature thresholds from 175 to 204, the edge strength value for the component containing the coldest temperature jumped from a value of 7.75 for a temperature threshold of 192 to 14.14 for a temperature threshold of 193. Between the temperature thresholds of 192 and 193, the connected component containing the coldest temperature was split from a component with 56 border points into two components, one containing 32 border points and the other 21 border points. The connected component for threshold 193 with 21 border points was the one which contained the coldest temperature and which represented that portion of the larger connected component obtained for threshold 192 around which there was a sharp edge. Considering the picture in its entirety, it was very doubtful if all points with infrared readings colder than 193 would have been more representative of the cumulonimbus portion of the sample than all points with infrared readings colder than 192. This sample was also chosen to illustrate one of the major problems of using border edge strength to

define cloud-type objects. Many cloud components (such as cumulonimbus) had a sharp temperature gradient on only one particular portion of the border of the component with very weak temperature gradients between, for example, the cumulonimbus portion and the anvil portion. Values for the border edge strength feature for a temperature threshold which contoured a cumulonimbus component of a sample window of satellite data would be considerably weakened by the presence of weak gradients between the borders of the cumulonimbus and cirrus portions of the sample.

Looking at the cold temperature end of the scale in Figure 3, one can see an increase in component edge strength value between infrared readings 187 and 188 and in total edge strength between 192, 193, and 194. Between the temperature thresholds of 187 and 188, the component containing the coldest temperature was sharply split. The connected component with the coldest temperature for threshold 187 had 276 border points whereas, after the split, the connected component with the coldest temperature for threshold 188 had 88 border points, meaning that two totally different cloud objects were being compared. The increase in total edge strength between 192, 193, and 194, however, seemed to correspond to an appropriate threshold for contouring the cumulonimbus portion of the sample. It was hoped that more sophisticated edge operators for calculating total border edge strength could find temperature thresholds for contouring the cumulonimbus portions of

Samples Nos. 1 and 2 (shown in Figures 1 and 2, respectively).

For the "low" cloud samples, Figures 5-7 and 12, only minor differences between edge strength values for the component containing the coldest temperature and total edge strength values can be found. Figure 5, representing middle clouds, is interesting because its almost perfect bell-shape suggests that the middle cloud portion of the sample can be isolated by selecting a threshold value near the peak of the bell-shaped curve, i.e., about 83. Figures 6, 7, and 12 reveal that, in general, for "low" clouds one expects low edge strength values. Higher edge values occurred primarily only for temperature thresholds where the number of border points around connected components was fewer than ten (see Figure 6, temperature thresholds 81 to 85). The statistical significance of a border edge strength feature based on such a small number of points is doubtful.

In sharp contrast to the low edge strength values for "low" clouds, the "mix" cloud samples, Figures 4 and 8-11, reveal high edge values with a break in the pattern of total border edge strength near an infrared reading of 100 (corresponding to a temperature of about 280°K). In Figure 4, the upward slope of the edge strength curve for the "low" cloud portion (left-hand side of the graph) seems to begin to level off near a temperature threshold of 100. The same break can be seen in

Figures 8, 10, and 11 and to a lesser extent in Figure 9 The "mix" samples in Figures 8-11 maintained, in general, high edge strength values throughout the entire infrared range above 100 in contrast to the cumulonimbus samples of Figures 1-3 for which edge strength values decreased for high (cold) infrared readings. In Figure 4, there was a slight decrease in edge strength values for temperature thresholds 175 to 180. However, edge values for the coldest temperature thresholds for the "mix" sample of Figure 4 did not drop as low as edge values for the coldest temperature thresholds of the "cumulonimbus" samples of Figures 1-3.

In conclusion, total edge strength values for all border points seemed to better discern thresholds for separating cumulonimbus portions of sample windows than edge strength values of particular components. In some cases, changes in the pattern of the total border edge strength feature with increasing infrared threshold corresponded to temperatures which appeared to physically contour distinct cloud-type regions. In other cases, no information could be obtained from the graphs of total border edge strength feature for threshold selection. The failure of the total border edge strength feature to segment these sample windows could be attributed either to a definition of edge strength which was not appropriate for meteorological data or to the inability of border edge strength features, in general, to contour cloud objects. The edge strength operators described in the next section were de-

signed to investigate the feasibility of using a total border edge strength feature to select appropriate thresholds for cloud-type objects.

4. Border Edge Strength Analysis

For each of the three edge operators designed for the experiments in this section, edge strength, EDGE(P), at a point P was defined as a function of the infrared temperature values of P and of its four vertical and horizontal neighbors. If the four neighbors of P are given by

```
  A
B P C
  D
```

and the infrared temperatures for points A, B, C, D, and P are given respectively by A_T, B_T, C_T, D_T, and P_T, then the edge strength at point P for each of the three edge operators is given by:

1) Minimum Edge Operator

$$EDGE(P) = P_T - \text{minimum}\ (A_T, B_T, C_T, D_T)$$

2) Average Edge Operator

$$EDGE(P) = P_T - \text{average}\ (A_T, B_T, C_T, D_T)$$

3) Directional Edge Operator

$$EDGE(P) = P_T - \text{minimum}\ (\text{average}(A_T, D_T), \text{average}(B_T, C_T))$$

Major differences between the three edge operators can be seen by comparing the edge strength values determined by each of the three edge operators on sample cloud-type objects. Let us first consider the sample object below

—	134	—
130	134	138
—	134	—

which illustrates a ramp edge in the vertical direction. The edge strength values for the three edge operators at

the center point are respectively 4, 0, and 0. From this example, one can see that the minimum edge operator has high values for a situation in which there is either change within a cloud pattern or a ramp edge between two cloud patterns. The average edge operator and the directional edge operator detect step edges such as the step edge in the vertical direction illustrated below

—	134	—
130	134	134
—	134	—

In this case, the edge strength values at the center point for the minimum edge operator, average edge operator, and directional edge operator are 4, 1, and 2, respectively. Note that the edge strength values calculated by the minimum edge operator were the same for both examples. One would expect high edge values for the minimum edge operator along both step edges and ramp edges between cloud patterns and also within cloud patterns in which there were large temperature variations. The average edge operator and the directional edge operator have very low values for ramp edges and high values for step edges. Edge strength values for the average edge operator for situations in which there is either a vertical or horizontal edge tend to be diluted by the temperature values of the horizontal or vertical neighbors, respectively. Note that for the step edge example, the edge strength value for the directional operator was higher than the edge strength value for the average

edge operator.

The edge strength value for a given temperature threshold T_i was defined as the average edge strength over all points P such that the infrared temperature reading of P was greater than T_i and at least one of the infrared temperature readings A_T, B_T, C_T, or D_T of the horizontal and vertical neighbors of P was less than or equal to T_i. This definition corresponds to the definition of the total edge strength feature presented in the previous section. The edge strength values for each of the three edge operators defined in this section can be compared with the edge strength values calculated from the Roberts gradient which appeared as dotted line graphs in Figures 1-12. Edge strength values obtained by applying the minimum edge operator, average edge operator, and directional edge operator to Sample Numbers 1-12 are plotted in Figures 13-24, Figures 25-36, and Figures 37-48, respectively.

Edge strength values for the first sample, which was classified as "cumulonimbus", can be found in Figures 1, 13, 25, and 37. Both the Roberts gradient technique (Figure 1) and the minimum edge technique (Figure 13) result in high edge values within cloud patterns with rapidly varying temperature profiles. In Figures 1 and 13, low edge strength values occur only for sea surface and/or low cloud portions and for the dense cumulonimbus portions. The cumulonimbus portion could be assumed to correspond to a pattern of relatively uniform or decreasing

edge strength values at the cold end of the infrared scale. Looking at Figures 1 and 13, one would expect that a good temperature threshold for the cumulonimbus portion might be either 195, 190, 185, or 180. From Figures 25 and 37, representing average edge strength values and directional edge strength values respectively, the best threshold for contouring a uniform block of cold temperatures appears to be 190. The entrance from a temperature-varying transition zone into a cloud object with relatively uniform temperatures can be detected at a lower threshold value by average or directional edge strength operators which are functions at any given border point of an object of points which are both interior and exterior to the object.

From an analysis of the graphs in Figures 2, 14, 26, and 38 for Sample Number 2, the threshold for contouring the cumulonimbus portion might be selected at either 185, 190, or 195. Referring to the raw picture data and the component analysis, one can determine that the infrared readings between 185 and 189 are not necessarily associated with the compact cumulonimbus portion of the sample but appear to be streaming off from the cumulonimbus portion and are probably part of the anvil. The temperature threshold of 190 was probably the best of the above three choices for segmentation of the cumulonimbus portion.

The choice of threshold value for the cumulonimbus portion of Sample Number 3 could not be exactly determined either from the graphs of Figures 3, 15, 27, and

39, or from the raw picture data. The best threshold seemed to be between 190 and 195. The change in threshold from 190 to 195 produced only minor differences in boundary of the largest connected component. The boundaries of many cloud-type objects are blurred because each infrared reading (from the instantaneous field of view of the satellite sensor) represents a weighted average of cloud-type blocks and "no cloud" blocks. The difficulty in determining an actual cloud object boundary for the cumulonimbus portion could be attributed both to this "smearing" effect and to the fact that the edge strength operators could not adequately distinguish transition zones between two cloud patterns from variable-temperature cloud objects. Whenever the operators saw varying temperature profiles, they treated them as instances of edges.

The edge strength values of "mix" Sample Number 4, which are plotted in Figures 4, 16, 28, and 40, failed to show a sudden downturn at the cold end of the temperature scale as was seen for the previous three "cumulonimbus" samples. In both Figures 28 and 40, there is very little change in magnitude between the edge strength values around 155 or 160 and those at the cold end of the scale. The edge strength values at the end of the scale in Sample Number 4 are larger than those for the cumulonimbus portions of Sample Numbers 1-3.

Sample Number 5, with edge strength values indicated on Figures 5, 17, 29, and 41, is the best example in the pilot study of the use of a peak in edge strength

values to locate a boundary of a cloud object. The peak of the curves (which occurs at approximately a threshold of 90) corresponds to the mid-point of a transition zone or ramp edge between a block of relatively uniform sea surface data probably mixed with some lower cloud and a block of relatively uniform middle cloud data. In all four previous examples, the peaks in edge strength values corresponded to temperatures within cloud objects with rapidly-varying temperature profiles.

The graphs for Sample Numbers 6 and 7, which are given respectively by Figures 6, 18, 30, and 42 and by Figures 7, 19, 31, and 43, illustrate that low clouds in general have low edge strength values. Edge strength values for low clouds calculated on visible data would probably have furnished a more definitive indication of a threshold between sea surface and low cloud data than can be found from these graphs. A temperature threshold of 75 seems to be the approximate point which separates cloud data from "no cloud" data for these two samples. Note that before a threshold of 75 (see Figures 18 and 19), the edge strength values showed a gradual upturn, probably indicating a change to a variable-temperature cloud object from a uniform sea surface.

The edge strength values for the "mix" samples, Numbers 8-11, are given in Figures 8, 20, 32, and 44, Figures 9, 21, 33, and 45, Figures 10, 22, 34, and 36, and Figures 11, 23, 35, and 47. The edge strength values are higher than those for

samples in the "low" and "cumulonimbus" classes and do not significantly taper off at the cold end of the temperature scale. The directional edge strength values are higher than the corresponding average edge strength values. This is, however, not the result of an edge in a particular direction but instead reflects a greater variability in the vertical direction than in the horizontal direction within a cloud-type pattern. The difference in directional variability can be attributed primarily to the difference between the vertical resolution of 4 miles and the horizontal resolution of 2 miles.

Sample Number 12, a "low" cloud sample, offers an interesting insight into a comparison of the four edge techniques for contouring a cloud-type object. The Roberts gradient edge strength values (Figure 12) peak at thresholds of 73 and 76 respectively, suggesting that an appropriate threshold for separation of sea surface and low cloud data would lie in this interval. The edge strength values for the average edge operator and the directional edge operator, given in Figures 36 and 48 respectively, peaked at the much colder reading of 84. The digitized cloud data in Sample Number 12 consisted of one connected component shaped approximately as shown below

```
XXXXXXXXXX
XXXXXXXXXX
XXXXXXXXXX
   XXX
XXXXXXXXXX
XXXXXXXXXX
XXXXXXXXXX
```

with sea surface temperatures relatively uniform and the temperatures within the cloud highly variable especially along the vertical direction. For the border points, the Roberts gradient and the minimum edge operator correctly detected a difference between the points and the exterior sea surface temperatures. The directional edge operator, however, looked primarily at the difference between the border point and the average of its horizontal neighbors which tended to be within the component. The average of the horizontal neighbors was lower than the average of the vertical neighbors because of the extreme variability within the cloud-type object. For this example, the peaks produced by the Roberts gradient and minimum edge strength operator corresponded to appropriate thresholds for contouring the cloud data.

The preceding analysis of the pilot samples in this study shows that the selection of thresholds to isolate physically significant temperature intervals based on edge strength values required human judgment coupled with feedback between edge strength values and the digitized meteorological satellite data. No particular edge strength operator seemed consistently superior to others. In general, the directional operator seemed best for separation of a cloud object of relatively uniform temperatures from a temperature-varying transition zone (for example, separation of cumulonimbus from cirrus by observation of a downturn in edge strength values or uniformly low edge strength values at the cold end of the temperature scale).

The minimum edge operator seemed best for separation of a cloud object containing rapidly-varying temperatures from a uniform background (for example, low cloud from sea surface by observation of the peak on an upturn in edge strength values at the warm end of the temperature scale). A threshold for separation of low cloud data from cirrus cloud data could not be easily selected from the graphs of edge strength values for any of the four techniques.

Selection of temperature contours which represent cloud-type objects need not be based on the spatial distribution of the infrared readings but can be approached from an analysis of the infrared histogram. The next section reports on the results of application of a cluster technique due to W. D. Fisher to the infrared histogram to determine clusters of temperature intervals corresponding to cloud-type layers.

5. Histogram Cluster Analysis

The most significant difference between the W.D. Fisher algorithm for clustering ordered sets of data points and other typical clustering algorithms is that "globally" optimal partitions of the data are obtained by dynamic programming procedures, rather than "locally" optimal partitions by iterative optimization procedures. The application of dynamic programming procedures in a clustering algorithm is feasible only when sufficient mathematical constraints are introduced to significantly reduce the number of possible partitions which must be examined at each step of the clustering algorithm. The number of partitions of M objects into K clusters is given by $\frac{1}{K!}\sum_{i=1}^{K}\binom{K}{i}(-1)^{K-i}i^{M}$ which is approximately $\frac{K^M}{K!}$. This number can be reduced from $O(K^M)$ to $O(M^K)$ by requiring each cluster to correspond to an interval of temperature values and can be further reduced by selection of an error criterion that is additive over clusters. Since an error function is calculated for each allowable partition, no initial starting values for cluster membership are required by the Fisher algorithm.

The form of the error function and the maximum number K_{MAX} of clusters desired must be specified by the user of the Fisher algorithm. The programs used in this section represent adaptations of basic programs for the Fisher algorithm which appear in Hartigan [3] on pp. 141-142. The error function for a partition of M objects into K clusters

was defined in terms of the cluster sum of squared deviations diameter. Let P(M,K) denote the partition of M ordered objects 1,2,...,M into K clusters given by

$$(I_1, I_1+1, \ldots, I_2-1),\ (I_2, I_2+1, \ldots, I_3-1), \ldots, (I_K, I_K+1, \ldots, M)$$

where $I_1 = 1$ and $I_{K+1} = M+1$, and let D(I,J), $1 \le I \le J \le M$, denote the cluster diameter given by

$$D(I,J) = \sum_{L=I}^{J} |X(L) - \bar{X}|^2$$

where $\bar{X} = \dfrac{\sum_{L=I}^{J} X(L)}{J-I+1}$

Then the error e[P(M,K)] associated with the partition P(M,K) is given by

$$e[P(M,K)] = \sum_{J=1}^{K} D(I_J, I_{J+1}-1).$$

There are four basic steps in the algorithm. The first step is to compute the diameters D(I,J) for all I,J such that $1 \le I \le J \le M$. Note that D(I,I) = 0 for all I. The second step is to compute the errors e[P(I,2)] for all paritions of I objects, $1 \le I \le M$ into 2 clusters. The optimal partition $\hat{P}(I,2)$ of I objects into 2 clusters is the partition which minimizes over all J, $2 \le J \le I$, the sum

$$D(1,J-1) + D(J,I).$$

Store in a matrix B of size MxK_{MAX}, for each value of I,

the lower boundary of the second cluster for the optimal partition $\hat{P}(I,2)$ as $B(I,2)$ and save the value of $e[\hat{P}(I,2)]$. For the third step, the errors for the optimal partitions of I objects into K clusters where $3 \leq K \leq K_{MAX}$ can then be obtained by finding the partition which minimizes over all J, $2 \leq J \leq I$, the quantity

$$e[\hat{P}(J-1,K-1)] + D(J,I).$$

Store the value of J which corresponds to the lower boundary of the Kth cluster of each optimal partition $\hat{P}(I,K)$ in the matrix B as $B(I,K)$. Then for step 4, one can discover the optimal partition of M objects into K clusters $\hat{P}(M,K)$ for any value of K, $K \leq K_{MAX}$, by backtracking as follows. The Kth cluster consists of the objects ranging from $B(M,K)$ to M. The (K-1)st cluster consists of objects ranging from $B(B(M,K)-1, K-1)$ to $B(M,K) - 1$, etc.

The definition of cluster diameter as a sum of squared deviations arises from the assumption that each of the clusters consists of independent observations drawn from a normal density. The statistical model for the partition P(M,K) into the K clusters

$$(I_1, I_1+1, \ldots, I_2-1),\ (I_2, I_2+1, \ldots, I_3-1), \ldots, (I_K, I_K+1, \ldots, M)$$

is that the observations $X(1), X(2), \ldots, X(I_2-1)$ are independent observations from the normal density $f(X|\theta_1) = \exp(-\frac{1}{2}(X-\theta_1)^2)/\sqrt{2\pi}$; the observations $X(I_2), X(I_2+1), \ldots, X(I_3-1)$ are independent observations from the normal density $f(X|\theta_2) = \exp(-\frac{1}{2}(X-\theta_2)^2)/\sqrt{2\pi}$; and so on, up to

$X(I_K)$, $X(I_K+1),\ldots,X(M)$ which are assumed to be independent observations from the normal density $f(X|\sigma_K) = \exp(-\frac{1}{2}(X-\sigma_K)^2)/\sqrt{2\pi}$. If the cluster diameter $D(I,J)$ is defined as above, then the sum of the cluster diameters is the same as the negative of the maximum log likelihood of the observations. Minimizing this sum (by use of an additive error criterion) means that those clusters have been found which make the given observations most probable. Detailed discussions of maximum likelihood estimation and relationships between various density functions and cluster diameters can be found in Duda and Hart [4] and Hartigan [3].

Theoretical models which apply to determination of the number of modes or clusters in the data are also presented in Hartigan [3]. If one assumes that the observations are normal and that the K+1 clusters of the optimal partition $\hat{P}(M,K+1)$ are obtained by splitting one of the K clusters of $\hat{P}(M,K)$ in two (which, however, usually is not the case), then the mean square ratio MSQ where

$$MSQ = (M-K-1)\left(\frac{e[\hat{P}(M,K)]}{e[\hat{P}(M,K+1)]}\right) - 1$$

is distributed as $F_{1,M-K-1}$. A large value of the mean square ratio means that K+1 clusters are necessary. A graph of the errors of optimal K partitions against values of K can also be used to select the best value of K. The correct number of clusters K is that number for which the decrease in error from the optimal (K-1) - partition to

the optimal K-partition is most noticeable.

The selection of the parameter K_{MAX} and the selection of the value of K for isolation of a cloud-type object consisting of the coldest temperatures in each sample of meteorological satellite data were based primarily on convergence properties of the values in the last (coldest) cluster as the number of clusters increased, on expected size of the standard deviation for cumulonimbus clusters, and on comparison of the objects isolated by the clustering algorithm for various values of K with stereoscopic views of the samples at two consecutive time periods, rather than on analysis of the sum of squares (error function) or the mean square ratio. Sharp decreases in the values of the sum of squares generally occurred as the number of clusters increased from 2 to 3 or 3 to 4 (primarily for cumulonimbus samples). No local peak was observed in values of the mean square ratio which would indicate that a particular value of K was to be preferred. Both these measures proved of little value in locating the mode of the narrow temperature range at the cold end of the scale which corresponded to a cumulonimbus cloud object.

It was found, experimentally, that in order to isolate the cumulonimbus portion of the "cumulonimbus" samples, a value of at least 10 had to be selected for K_{MAX}. The parameter K_{MAX} can be compared to the cluster size parameters of those clustering algorithms for which the number of desired clusters is not specified by the user. For example, in the discussion of the ISODATA routine of Endlich

et al. [5], the maximum allowable radius of a cluster was a function of a parameter called the sphere factor which was selected as 0.7 to yield 10-15 clusters per grid area. If the splitting and lumping options of the ISODATA routine were used, one or more additional parameter values had to be chosen by the user. For the Fisher algorithm, a large value of the parameter K_{MAX} was necessary in order to isolate clusters consisting of a narrow range of cold temperature values. Large values of K tended to break both the low cloud portion and the cirrus cloud portion of "cumulonimbus" samples into many clusters. The criteria used in this study for merging these clusters were based on expected temperature signatures. Various other criteria based on cloud-type characteristics need to be examined.

Although optimal partitions were obtained for each value of K from 1 to 10, a value of K=9 was selected as that value which best described the cumulonimbus portion of "cumulonimbus" samples. The infrared histograms for each of the twelve samples in the pilot study are shown in Figures 49-60. The lower boundary of the ninth cluster (for K=9) is marked on each of the histograms. Descriptions of the coldest of nine clusters for each of the pilot samples appear in Table 2. From the entries in the column of partition sizes, one can observe that, for each sample in the pilot study, as the value of K was increased, there was a sequence of optimal partitions of length at least 2 which yielded the same coldest cluster. This sequence included the value of K=9 for every sample. This

type of stability or convergence of the observations which comprised the coldest of nine clusters did not occur, in general, in the case of "cumulonimbus" and "mix" samples, for clusters other than the coldest cluster for any values of K which were examined.

The statistical properties of the coldest of nine clusters for "cumulonimbus" samples were of the order of magnitude that would be expected for cloud objects consisting only of "cumulonimbus" data points. For example, the cloud-type signatures for cumulonimbus samples derived by Greaves and Chang [6] for Nimbus 2 data consisted of a mean temperature of 224°K and a standard deviation of 6.4°K. Converting the mean and standard deviation of infrared observations into temperature values by use of the calibration table found in Table 1 of [1], one can see that the mean temperatures of the coldest of nine clusters for "cumulonimbus" samples 1-3 ranged from approximately 213°K to 222°K with standard deviations approximately 5°K. Choosing a cluster size for the three "cumulonimbus" samples in the pilot study which decreased the lower boundary of the coldest of nine clusters for K=9 by more than two counts (infrared observations) resulted in standard deviations of 6.5, 7.5, and 8.2, respectively, which seemed too large. Note from Table 2 that the value of the standard deviation for the ninth cluster varied with cloud type. For "low" cloud samples 5-7 and 12, the standard deviation values ranged from 1.51 to 1.95. For "cumulonimbus" samples 1-3, the standard deviations ranged from 4.01 to

4.32. Higher values, ranging from 4.67 to 7.14, characterized the coldest of nine clusters of "mix" samples. The cluster standard deviation is correlated to the cluster edge per unit area feature which was successfully used to classify cluster cloud type (see Section 6).

The third consideration which prompted a choice of K=9 for the number of clusters to isolate the cumulonimbus portion of "cumulonimbus" samples was a comparison between the outlines of the coldest cluster for various values of $K \leq 10$ with the stereoscopic projection of the two consecutive infrared satellite photographs of the original data. The pictures of the data (reproduced in Figures 2 and 4 of [1]) were displayed under a mirror stereoscope and aligned by hand in the approximate direction of upper level wind flow until a position was reached which seemed to best segregate layers of cloud motion. The outlines of the coldest clusters for values of K=9 approximately coincided with outlines of cumulonimbus layers which lay on the same level (as seen under the mirror stereoscope). Pictures of "cumulonimbus" samples 1-3 and outlines of the cumulonimbus portion of the samples which were obtained by selecting the coldest of nine clusters are included in Section 6 (see Figure 64). The points with infrared readings in the range delineated by the coldest cluster for K=9 appear as solid white in the segmented pictures of the "cumulonimbus" samples.

Another interesting comparison can be made between the range of observations included in the coldest clusters for

"cumulonimbus" samples 1-3 and the pattern of edge strength values at the right-hand side of the graphs of Figures 37-39. For "cumulonimbus" samples 1 and 2, both the cluster analysis and the edge strength analysis indicate a change in cloud-type pattern occurring near a threshold of 190. For Sample Number 3, the edge strength analysis indicates the possibility of a higher threshold -- perhaps about 195 -- and the cluster analysis indicates a change in pattern at about 198 for values of K from 7 to 10 and 196 for values of K from 5 to 6. A threshold lower than 195 produced small regions which did not seem to physically relate to the main cumulonimbus cloud component observed in the picture data.

Since no segmentation of "low" cloud samples was desired, it was not necessary to apply the more expensive cluster analysis procedures to "low" cloud samples if they could be identified by statistical classification techniques. For example, from Experiment Number 2 of Table 30 in [7], we can see that maximum likelihood classification on the design set using only one feature -- coldest temperature value in the sample -- correctly identified 95.4% of the "low" cloud samples. An efficient procedure for "low" cloud classification would be to utilize temperature thresholds derived either from labelled samples or from knowledge of the three-dimensional temperature structure of the atmosphere at the given geographical location to isolate "low" clouds. For the samples in the pilot study, it was decided to classify all samples for which the

maximum infrared reading was less than or equal to 100 (i.e., coldest temperature in the sample was $\geq$ 280°K) as "low" cloud samples.

The coldest temperature feature can also be used to identify "mix" cloud samples containing dense cirrus in which the coldest temperature is too warm for "cumulonimbus" samples. The percentage of "mix" samples correctly classified using the coldest temperature feature was approximately 71% in Experiment Number 2, Table 30 of [7]. A range of temperatures for a given geographical area can be specified for isolating "mix" samples. For the experiments in this section, all samples for which the coldest infrared reading in the sample fell within the range from 150 to 170 (i.e., coldest temperature value in the sample lay between 255°K and 245°K) were classified as "mix". Note from Table 9 of [7] that for the coldest temperature feature (Number 303), the mean value for "cumulonimbus" samples was approximately 46 which corresponds to a temperature of 206°K (160+46) with a standard deviation of approximately 19°K. It is very unlikely that the coldest temperature values for "cumulonimbus" samples would lie in the range between 245°K-255°K for latitudes between the equator and 12.5°N.

The decision procedure for classification of the samples in the pilot study is presented in Figure 61. At the first stage of the decision tree, all samples of meteorological satellite data in which the maximum-infrared reading is less than or equal to 100, i.e., coldest temperature $\geq$ 280°K, are classified as "low" clouds. The

Fisher clustering algorithm for a value of K=9 is then applied to all other samples. At the second stage of the decision tree, the samples are divided into samples which may be either "low" or "mix", "mix" samples, and samples which may be either "mix" or "cumulonimbus". The criterion used to separate the samples at this stage consists of determining two thresholds on the temperature line. For the experiments in this section and in the next chapter, threshold values of 150 (corresponding to approximately 255°K) and 170 (corresponding to approximately 245°K) were selected. Samples for which the maximum infrared reading lay between 100 and 150 were tested at the third stage of the decision procedure to determine if they were "low" or "mix" samples. Samples for which the maximum infrared reading was greater than 170 were tested at the third stage of the decision procedure to determine if they were "mix" or "cumulonimbus" samples. Samples for which the maximum infrared reading lay between 150 and 170 were classified as "mix" samples.

The tests applied at the third stage of the decision procedure consisted of calculating feature values for the coldest of the nine clusters and using these feature values to classify the sample. Various different features could be applied. Two features that were examined were a visible brightness feature and an infrared cluster edge strength feature. One would expect from previous experiments that the visible brightness feature would be appropriate for separation of "cumulonimbus" and "mix" samples

but might not be as useful for separation of "low" and "mix" samples (see Fisher distance values for Feature Number 113 in Table 18 of [7]). The cluster edge strength feature, which is correlated to the standard deviation and mean texture features for the diagonal directions (see, for example, Fisher distance values for Feature Number 302 of Table 19 of [7] and Feature Numbers 323 and 324 of Table 24 of [7]), would be expected to separate both "low" from "mix" samples and "mix" from "cumulonimbus" samples.

At the fourth stage of the decision procedure, clusters were merged together using a temperature threshold value to determine low cloud segments and cirrus cloud segments. The cumulonimbus cloud segment of samples which were classified as "cumulonimbus" at the third stage of the decision procedure consisted of the coldest cluster. The low cloud segment was obtained by merging together all clusters for which the minimum infrared observation within the cluster was less than or equal to 100. The cirrus cloud segment was obtained by merging together all clusters (except the coldest cluster for "cumulonimbus" samples) for which the minimum infrared observation within the cluster was greater than 100.

Two different algorithms for classification of cloud samples at the third stage of the decision procedure (Figure 61) -- one based on a visible brightness feature and one based on a cluster edge strength feature -- are formulated in Figures 62 and 63, respectively. Classification in both of the algorithms results from comparisons

of the mean of a feature value calculated on points in the coldest cluster with one or more means of the same feature value calculated on points lying within a middle range of temperature values and points lying within a low range of temperature values. The temperature values of the points within the middle range suggest that the cloud type of these points is probably either cirrus or cirrus mixed with lower clouds. The temperature values of the points within the low range are indicative of low cloud types. Thus, the classification scheme of both algorithms depends on comparison of feature values of the coldest cluster with feature values of known cloud types within the same sample rather than on a threshold of feature values.

A threshold scheme could have been used for classification. For example, one could have said that if the edge strength feature of the coldest cluster for samples with a maximum infrared reading above 170 was greater than a given threshold value E_T, the sample would be classified as "mix" instead of "cumulonimbus"; or, if the visible brightness feature of the coldest cluster was greater than a given value V_T, the sample would be classified as "cumulonimbus" instead of "mix". Similar thresholds could have been determined for samples for which the maximum infrared reading lay between 100 and 150. For these samples, one would expect that the values of both the visible brightness feature and the cluster edge strength feature would be higher for "mix" samples than for "low" samples. One can deduce appropriate threshold values by looking at the

spread of feature values given in Table 3.

The comparison of feature values for the coldest cluster with feature values within the same sample for points whose infrared readings fell within the middle and/or low ranges was independent of choice of threshold value. This type of algorithm was pursued, rather than classification algorithms based on thresholds which would probably not have been valid for all samples over a wide geographical area. The middle and low ranges were selected so that the difference between the maximum and minimum infrared reading within each of these ranges would be the same as the difference between the maximum and minimum infrared reading within the coldest cluster. This criterion was particularly important for the computation of the mean cluster edge strength feature.

The mean cluster edge strength feature was defined in terms of the Roberts gradient values for each point. The value of the Roberts gradient at a point A in the upper left-hand corner of a 2x2 array of points $\begin{matrix} A & B \\ C & D \end{matrix}$ was defined as

$$\max(|A-D|, |C-B|).$$

The cluster edge strength feature for a set of points was defined as the mean value of the Roberts gradient for all points within the set.

The visible brightness feature of a set of points was defined as the mean value of the brightness reading for all points in the set for which visible data was available.

The Fisher clustering algorithm was applied to the first twelve samples of the large size infrared windows of Data Set III. The resolution of these infrared windows was 2 miles in the horizontal and 4 miles in the vertical. The visible windows corresponding to this set (i.e., same center point as infrared windows) had a resolution of 2 miles in both the horizontal and vertical directions. Consequently, for all data points lying in the upper fourth and lower fourth of each infrared window, there was no corresponding visible data. This means that the analysis using the visible brightness feature was limited to the small size 64x32 infrared windows of Data Set III.

A comparison of the classification results for "cumulonimbus" samples using both the visible brightness feature and the cluster edge strength feature can be made by referring to Table 3. Both algorithms correctly classified the first three samples as "cumulonimbus". The visible brightness feature for the coldest clusters of all three samples was greater than the visible brightness feature for the sets of points comprising the middle range sets. Values for the cluster edge strength feature of the coldest cluster differed to a greater extent from edge strength values for sets of points in the middle range than edge strength values for sets of points in the low range. For the "cumulonimbus" samples, cluster edge strength values for sets of points in the middle range were uniformly higher than cluster edge strength values for the coldest cluster and for sets of points in the low range.

The same decision procedure and classification algorithms that were applied to the "cumulonimbus" samples were applied to "mix" samples Numbers 4, 8, 9, and 10. The maximum infrared reading in these "mix" samples was greater than 170. The visible brightness feature algorithm incorrectly classified "mix" samples Numbers 4 and 10, as "cumulonimbus" whereas the cluster edge strength feature algorithm correctly identified them as "mix". Note that although the visible brightness features calculated on the coldest cluster of these two samples were higher than the visible brightness features calculated on sets of points in the middle range, the values for the coldest cluster for both samples were lower than corresponding values for the "cumulonimbus" samples. The difference was so slight, however, especially in the case of Sample Number 4, that one would hesitate to adopt a threshold method on the visible brightness feature values to distinguish "cumulonimbus" from "mix" samples. No visible data was available for "mix" Sample Number 9 since the coldest cluster lay entirely in the upper fourth of the large size infrared window. "Mix" Sample Number 8, in contrast to Samples Numbers 4 and 10, was correctly classified by the visible brightness feature algorithm. All the "mix" samples in the pilot study were correctly classified by the decision procedure of Figure 61 combined with the cluster edge strength feature algorithm of Figure 63. The cluster edge strength feature values for the coldest cluster were markedly similar to cluster edge strength feature values

for the set of points in the middle range -- indicating that the coldest cluster of "mix" samples should be merged into those clusters which were contained in the set of points comprising the middle range.

Classification of the "mix" Sample Number 11, and "low" samples Numbers 6, 7, and 12, was done on the basis of the value of the coldest temperature (maximum infrared reading) within the sample. Following the decision procedure of Figure 61, the "mix" Sample Number 11 was classified as "mix" since the maximum infrared reading in the sample (a value of 158) lay between the values of 150 and 170. "Low" samples Numbers 6, 7, and 12 were classified as "low" since the maximum infrared reading in the sample was less than 100.

The only sample in the pilot study which travelled down the branch of the decision procedure for samples with maximum infrared reading between 100 and 150 was "low" Sample Number 5. This sample consisted primarily of middle clouds. From the values of the visible brightness features and cluster edge strength features, one would guess that middle clouds are, in general, brighter and denser than lower clouds. No distinction was made in this study between middle clouds and low clouds. Since the major purpose of this study was to isolate temperature layers suitable for upper level wind estimates, middle clouds were grouped together with low clouds into a "low" class. The cluster edge strength feature algorithm classified Sample Number 5 as "low" since the coldest portion of the sample

exhibited the same dense effect (lack of edge variation within the cluster) as one would expect for low clouds. The visible brightness feature algorithm expected the coldest portion of a "mix" sample to be brighter than the coldest portion of a "low" sample, and thus classified Sample Number 5 as "mix".

The classification and segmentation performance of the Fisher clustering algorithm combined with the cluster edge strength feature algorithm was superior to that of all the segmentation techniques which were developed in this section and applied to the twelve samples in the pilot study. This combination of the decision procedure of Figure 61 with the algorithm of Figure 63 formed the cluster edge strength model. A more extensive test of this model was conducted on the large size infrared windows of Data Set III. The results of this test are presented in the next section.

6. Experimental Evaluation of Cluster Edge Strength Model

The cluster edge strength model for automatic cloud classification and segmentation was tested on the 107 large size infrared windows of Data Set III (described in [1]. Classification results for each sample are shown in Table 4. Pictures of the original and segmented sample windows appear in Figures 64-70. The cluster edge strength model resulted in cloud sample segmentation by cloud type (low, cirrus, and cumulonimbus) and cloud sample classification results which were consistently superior to those obtained by statistical classification techniques. The number of test samples correctly classified by the cluster edge strength model was 102. The maximum number of test samples (107 small size windows of Data Set III) correctly classified by any of the experiments in Chapter IV using statistical classification techniques and visible and infrared features was 87. The maximum number of test samples (107 large size windows of Data Set III) correctly classified using statistical classification techniques and infrared features only was 70. These comparative results demonstrate the superiority of the cluster edge strength model for automatic cloud classification and segmentation.

6.1 Classification Results

The 107 large size infrared test samples of Data Set III consisted of 44 samples of "low" cloud, 30 samples of "mix" cloud, and 33 samples of "cumulonimbus" cloud. The geographical region from which the test samples were selected contained a high percentage of upper-level cloud types; however, there was no test sample which contained only cirrus clouds. Within a test sample cirrus clouds occurred either in combination with cumulonimbus clouds (in which case the sample was labeled by meteorologists as "cumulonimbus") or in combination with lower clouds (in which case the sample was labeled as "mix").

Statistical classification techniques were used to classify 37 of the 107 samples on the basis of maximum (coldest) infrared reading within the sample. Since the maximum infrared reading of 33 of the samples was less than or equal to 100, these samples travelled down the left-hand side of the cluster edge strength model decision tree (Figure 61) at Stage 1 and were accordingly classified as "low". Cluster analysis techniques were applied to the remainder of the 74 samples. Four of these samples -- those with maximum infrared reading between 150 and 170 -- travelled down the central branch of the cluster edge strength model decision tree at Stage 2 and were classified as "mix".

At Stage 3 of the decision procedure, cluster edge strength features were calculated for each of the 26 test samples with maximum infrared reading between 100 and 150 and for each of the 44 test samples with maximum infrared

reading greater than 170. The procedure for defining temperature interval clusters within these samples and for classifying the samples by comparing edge strength feature values on the specified temperature interval clusters is summarized in Figure 63. The cluster edge strength feature values obtained using this algorithm can be found in Table 4. The cluster edge strength decision procedure classification for each sample and the meteorological (cloud-truth) classification are also presented in Table 4.

It can be seen from Table 4 that 5 of the 107 test samples -- Sample Numbers 19, 59, 73, 79, 91 -- were misclassified. For the first four samples, the misclassification seemed to result from the fact that sample points in the coldest cluster were few in number and located too close to the border to obtain a statistically representative edge strength feature value of the cloud pattern in which the coldest cluster was embedded. For Sample Numbers 19, 59, and 79, the cumulonimbus portions of the samples existed only either near the right-hand border of the sample (Sample Numbers 19 and 59) or near the top of the border (Sample Number 79). The coldest cluster in Sample Number 73 consisted only of a few points near the top of the border of the sample. Comparison of edge strength feature values between the coldest cluster and the low range cluster led to the misclassification of the sample as "mix" instead of "low". There was no obvious explanation for the misclassification of Sample Number 91.

6.2 Segmentation Results

The sample image data and sample segmentation results obtained from the cluster edge strength model decision procedure are shown in Figures 64-70. These figures contain the original and segmented images for Sample Numbers 1-16, 17-32, 33-48, 49-64, 65-80, 81-96, and 97-107, respectively. Images made from the original digitized test data appear on the left-hand side of each figure and the segmented images on the right-hand side of each figure. The sample test windows are arranged within each picture in row-major order. That is, in Figure 64, Sample Number 1 is in row 1, column 1; Sample Number 2 is in row 1, column 2;...; Sample Number 5 is in row 2, column 1;...; and Sample Number 16 is in row 4, column 4.

A re-scaling of the infrared gray scale had to be applied to the original digitized image data in order to create the pictures shown on the left-hand side of Figures 64-70. The minimum infrared reading within any of the 107 large size infrared test windows was 55 and the maximum infrared reading was 211. The picture output device could record a maximum of 64 gray levels. In order to retain maximum possible resolution for display of the upper end of the infrared scale (cold end), each infrared gray level value I was transformed into a corresponding integer output value G as follows:

$$G = 0, \text{ for } I < 92$$

$$G = \text{entier}\left(\frac{I-90}{2}\right), \text{ for } I \geq 92$$

where entier(x) ≡ greatest integer less than or equal to x.

Consequently, the warmer portions of low-level clouds appear uniformly dark in the picture display of Figures 64-70. The original digitized sample data can also be seen in picture form in Figure 2 of [1]. The latitude and longitude of the center point of each of the test samples is given in Table 3 of [1]. However, even when overlaying a latitude-longitude grid on the digitized sample data, it is difficult to exactly locate each sample window and, in particular, to judge the areal extent of each window. The problem is further complicated by differences in resolution between Figure 2 of [1] (4x4 n.mi. resolution) and the infrared digitized data (2x4 n.mi. resolution). The pictures in Figures 64-70 show the exact test sample windows on which the cluster edge strength model was evaluated. The pictures in [1] were invaluable aids for determining cloud-truth classification and for examining cloud layer motion differences.

In order to evaluate the segmentation results, two questions must be asked:

1) Did the segmentation results lead to correct classification results?
2) Do the segmentation results correspond to known cloud type and/or cloud motion layers within the sample?

The classification results were discussed in the previous section of this chapter and tabulated in Table 4. The

segmentation results appear in pictorial form in Figures 64-70.

Looking at the segmented sample windows on the right-hand side in Figures 64-70, the sample classification can be determined from the number of colors in the sample. If white, gray, and black segments appear in the sample, the sample was classified as "cumulonimbus". The white segments represent the cumulonimbus portions of the sample determined by the cluster edge strength model decision procedure, the gray segments are the cirrus portions, and the black segments are the low cloud portions. If there are only gray and black segments in the sample, the sample was classified as "mix". All sample windows which appear uniformly black were classified as "low". Five samples were misclassified -- Sample Number 19 in row 1, column 3 of Figure 65; Sample Number 59 in row 3, column 3 of Figure 67; Sample Number 73 in row 3, column 1 of Figure 68; Sample Number 79 in row 4, column 3 of Figure 68; and Sample Number 91 in row 3, column 3 of Figure 69. The misclassification of the first four samples is due to proximity of coldest cluster points to sample borders, as can be verified by examining the picture windows above.

Quantitative evaluation of the quality of automatic segmentation results requires precise "a priori" meteorological location of segmentation boundaries. Location of appropriate segmentation boundaries for cloud type and/or cloud motion layers in sample windows of meteorological satellite data often demands contextual and motion informa-

tion which is not visually available upon inspection of the original sample window data. According to Hubert [8], "the analyst decides on cloud type by first determining the synoptic situation... Once the synoptic situation is determined, observations of cloud characteristics and cloud motion aid the analyst in segregating cloud layers and in specifying cloud types." Information from many different sources (whether satellite- or ground-based) must be merged and geographically registered in order to outline segmentation boundaries. If the time-consuming and tedious procedure of manually outlining segmentation boundaries is to result in sufficiently precise boundaries for the quantitative assessment of the quality of automatic segmentation results, then further refinements in techniques for identifying and registering borders of cloud segments within sample windows with cloud motion borders deduced from analyses of animated cloud motion data must be developed.

Although a quantitative evaluation of the segmentation results was not feasible, a qualitative evaluation was possible and was used to direct the development of segmentation techniques. With the aid of a satellite meteorologist, information from digital computer output, from pictures of original sample windows in Figures 64-70, from visual and infrared photographs at two consecutive time periods (see Figures 1-4 of [11]) examined under a mirror stereoscope, and from film loops was used to assess the cloud-type and/or cloud motion layer configurations in individual samples.

7. Summary and Conclusions

Three segmentation models were developed. Each of the models assumed that cloud-type objects were characterized by non-overlapping temperature interval layers. The first model defined cloud-type objects as connected components with sharp gradients along the border. The second model characterized nonuniform cloud-type situations as multiple cloud layers with unique temperature structures separated by sharp edge gradients between borders of cloud layers and/or sea surface. The third model assumed that cloud layers would appear as clusters in infrared histograms.

Cloud-type thresholds in the first model were selected from an analysis of the change in average border edge strength of the connected component containing the maximum (coldest) infrared reading as the temperature contour of the component was varied. Average border edge strength was calculated using a Roberts gradient operator. The model proved too sensitive to artificial changes in average border edge strength resulting either from splitting of a connected component into two parts or from an insufficient number of border points for computation of an average border edge strength feature.

The second segmentation model differed from the first with respect to the method of calculation of average border edge strength feature and operators used to define the edge at border points. The average border edge strength feature for a particular threshold (temperature contour) was obtained by averaging the edge strength values over border

points of all connected components, not just the connected component containing the coldest infrared reading. Three edge operators -- a minimum edge operator, an average edge operator, and a directional edge operator -- were designed to detect changes between cloud layers. In general, the directional edge operator seemed best for separation of a cloud layer consisting of relatively uniform temperatures from a temperature-varying transition zone, and the minimum edge operator seemed best for separation of a cloud layer containing rapidly-varying temperatures from a homogeneous background. The problem of threshold selection on the basis of the average border edge strength feature was complicated by the fact that steep border gradients on one side of a cloud object were sometimes weakened by fuzzy gradients on another side, with the resultant average border edge strength feature being of the same magnitude as gradients existing within cloud objects.

The third model based segmentation decisions on an examination of clusters in the infrared histogram and features (visible or infrared) extracted from the clusters. Both a visible brightness feature and an edge strength per unit area feature were used to determine whether or not a cluster represented a distinct cloud object or should be merged with other clusters to form a cloud-type object. The incorporation of statistical pattern recognition techniques with this segmentation model to identify the cloud type of cloud objects formed the cluster edge strength model.

The cluster edge strength model combined statistical pattern recognition techniques with the segmentation techniques of the third segmentation model above. The radiative properties of cloud types were used to classify samples with typical "low" and "mix" temperature profiles. Specific thresholds could be derived either from radiative transfer models, from labeled design samples, or from experience. Cluster analysis to identify cloud-type segments was applied to all samples except those which were classified as "low" based on maximum (coldest) infrared reading. Edge strength per unit area averaged over all cloud data with temperatures in the interval defined by the coldest cluster was compared with edge strength per unit area for specific temperature intervals of comparable width. The comparison resulted in cloud-type classification and segmentation of windows of satellite data.

The cluster edge strength model demonstrated consistent superiority to statistical pattern recognition models for classification of meteorological satellite data both in terms of classification accuracy and practicality. The tedious preparation of a labeled set of design samples was not required to obtain decision boundaries. The cluster edge strength model used only infrared data (thus simulating night-time conditions), two features (maximum infrared reading and segment edge strength per unit area), and a computationally efficient clustering technique to achieve a classification accuracy of 95%. Statistical pattern recog-

nition models achieved a maximum classification accuracy on independent data of 81% using both visual and infrared features and 65% using only infrared features.

The incorporation of image segmentation techniques into operational cloud classification systems for preprocessing meteorological satellite data is strongly recommended. The cluster edge strength model is one example of a model that merges both statistical pattern recognition techniques and image segmentation techniques to identify complex cloud-type conditions. Similar models could be developed for specification of cloud and surface conditions from meteorological satellite data in various spectral regions.

References

[1] J. A. Parikh, "Cloud classification: experimental evaluation", Computer Science Center, Univ. of Maryland, College Park, Technical Report 514, March 1977.

[2] A. Rosenfeld and A. C. Kak, Digital Picture Processing. New York: Academic Press, 1976.

[3] J. A. Hartigan, Clustering Algorithms. New York: John Wiley & Sons, 1975.

[4] R. O. Duda and P. E. Hart, Pattern Classification and Scene Analysis. New York: John Wiley & Sons, 1973.

[5] R. M. Endlich, D. E. Wolf, D. J. Hall, and A. E. Brain, "Use of a pattern recognition technique for determining cloud motions from sequences of satellite photographs", Journal of Applied Meteorology, vol. 10, no. 1, pp. 105-117, February 1971.

[6] J. R. Greaves and D. T. Chang, "Technique development to permit optimum use of satellite radiation data", Goddard Space Flight Center, Greenbelt, MD, Final Report on NASA Contract N62306-69-C-0227, May 1970.

[7] J. A. Parikh, "Cloud pattern classification from visible and infrared data", Computer Science Center, Univ. of Maryland, College Park, Technical Report 442, February 1976.

[8] L. F. Hubert, "Techniques for deriving winds from cloud movement", Proceedings of the Nineteenth Annual Meeting of the American Astronautical Society, Dallas, TX, June 19-21, 1973.

Figures 1-12

Comparison of border edge strength feature values of Samples Nos. 1-12 for coldest connected component (solid curve) and for all connected components (dotted curve) using Roberts gradient operator.

Figure 1

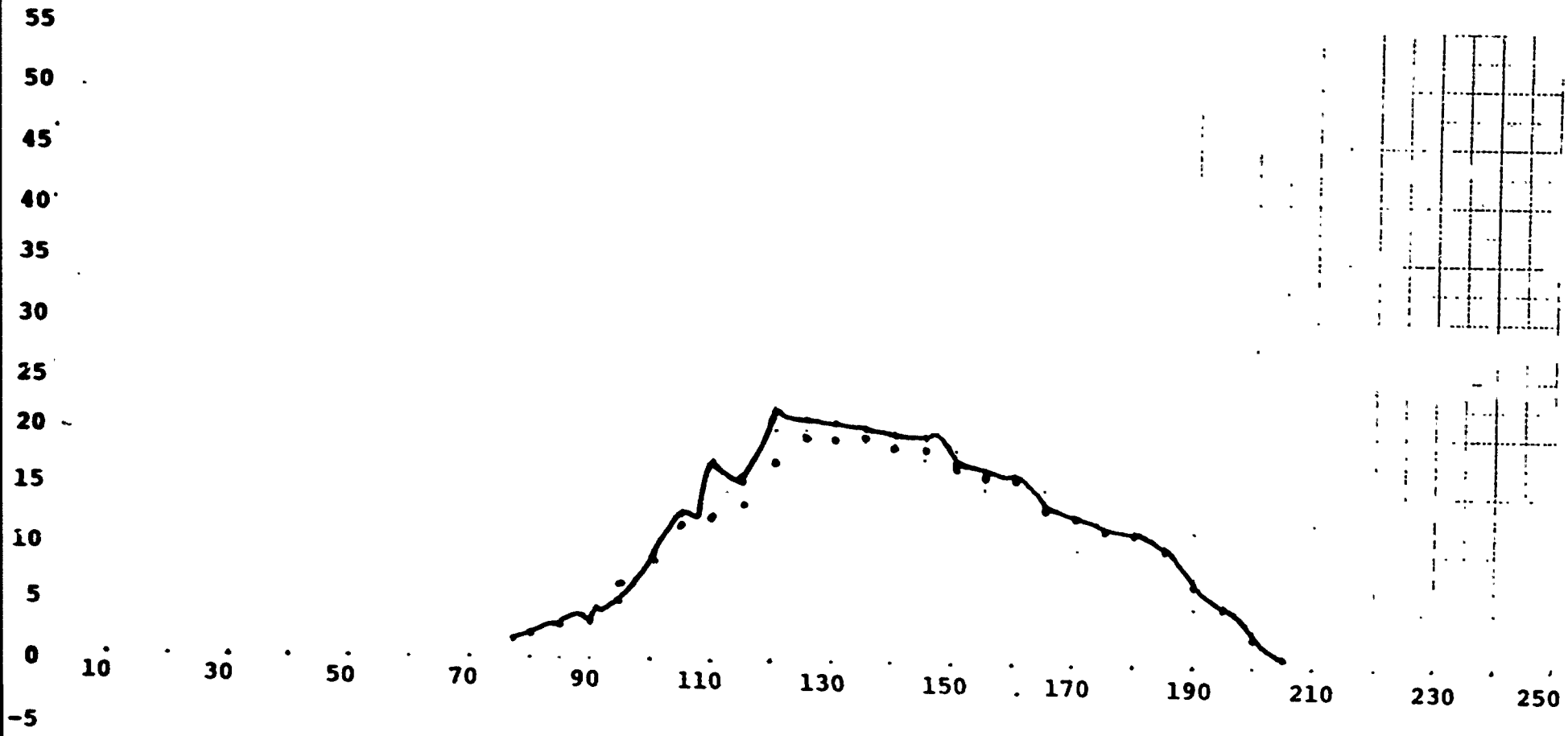

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 2

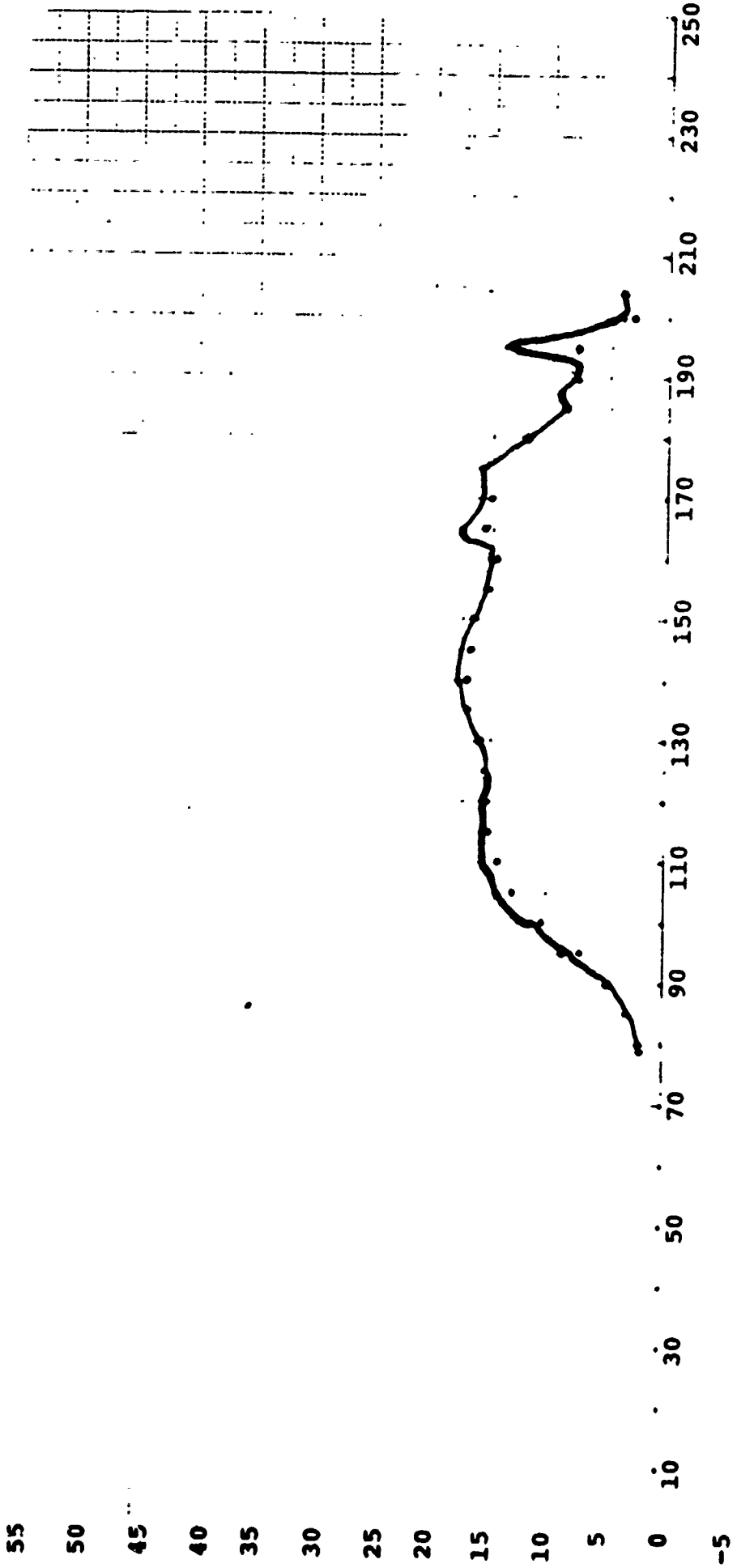

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figure 3

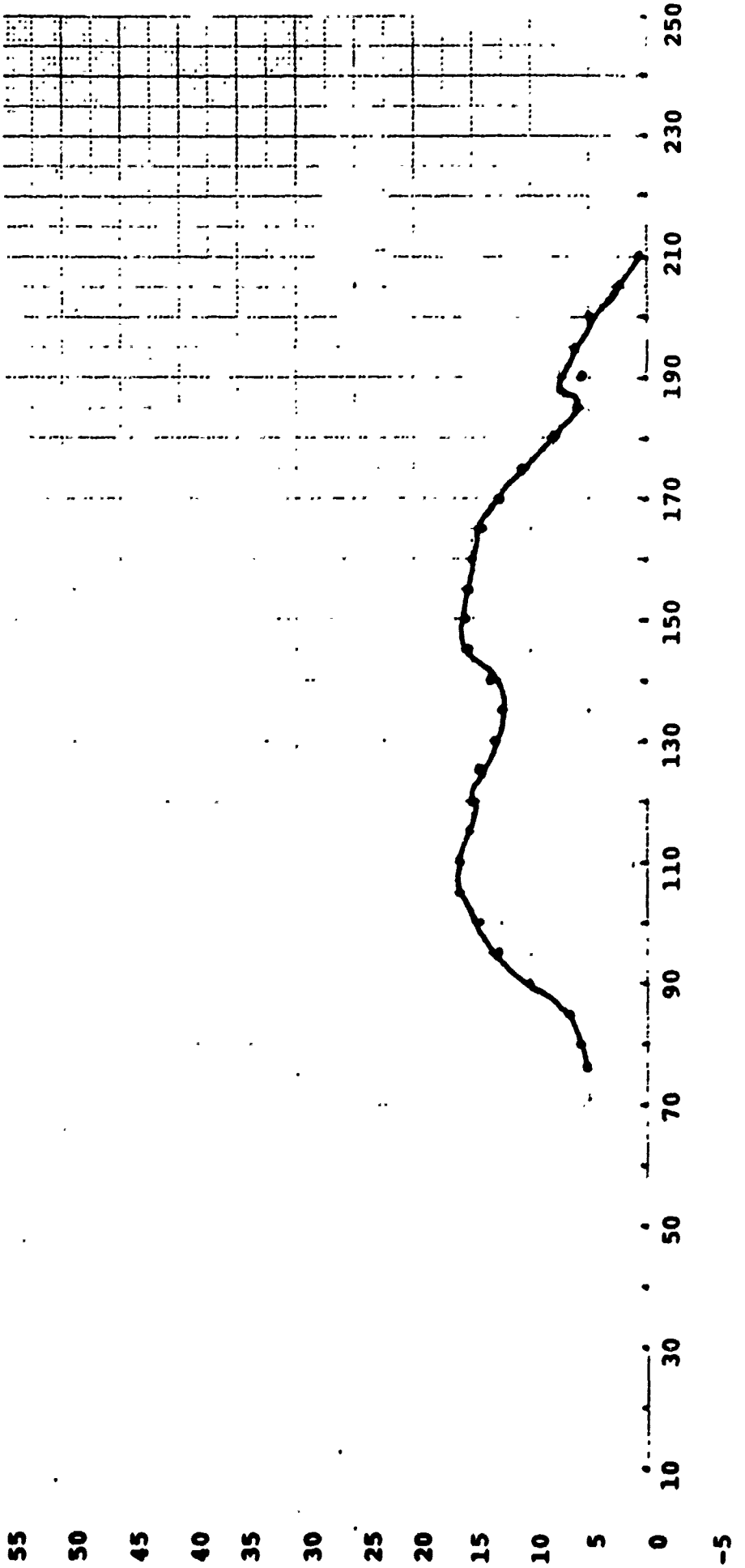

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figure 4

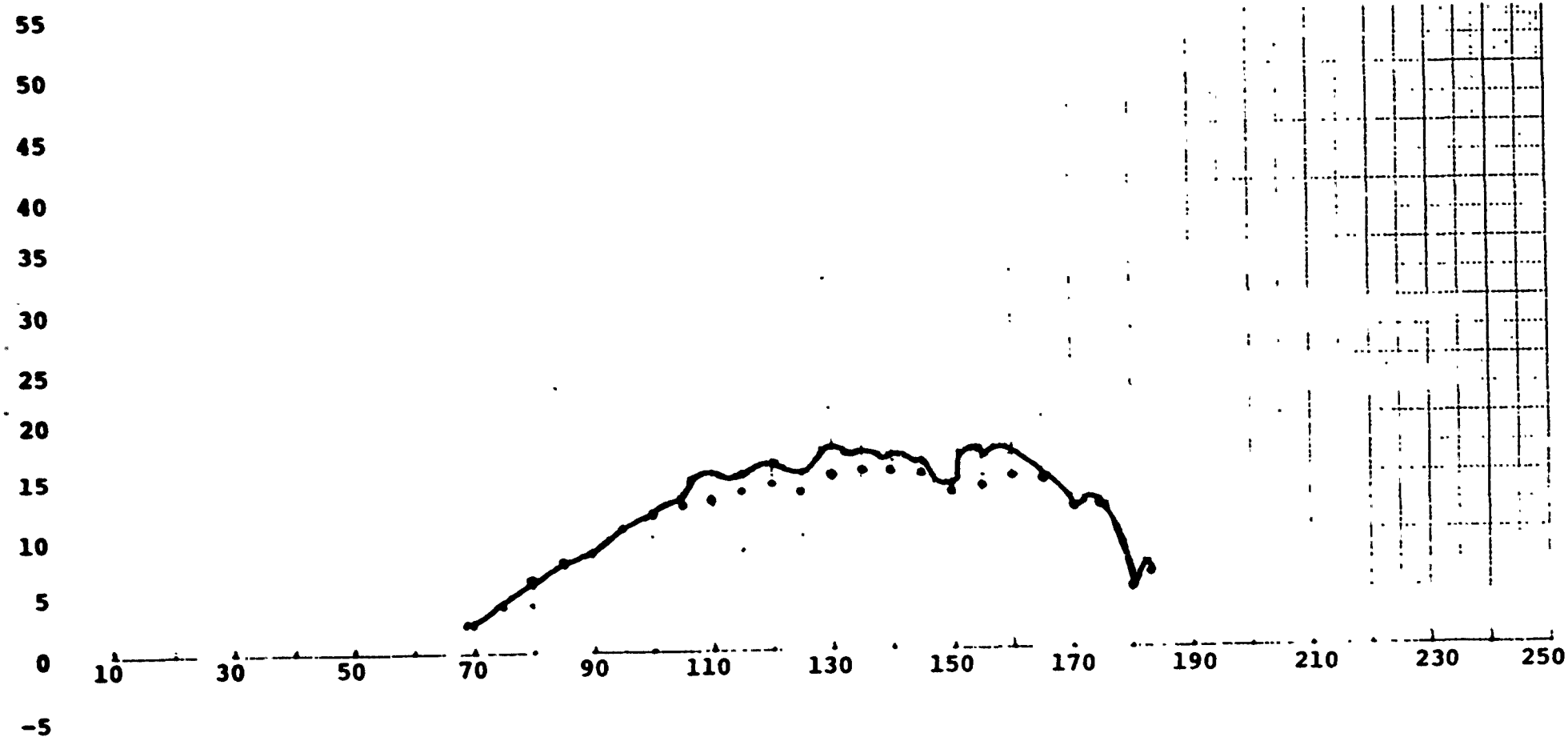
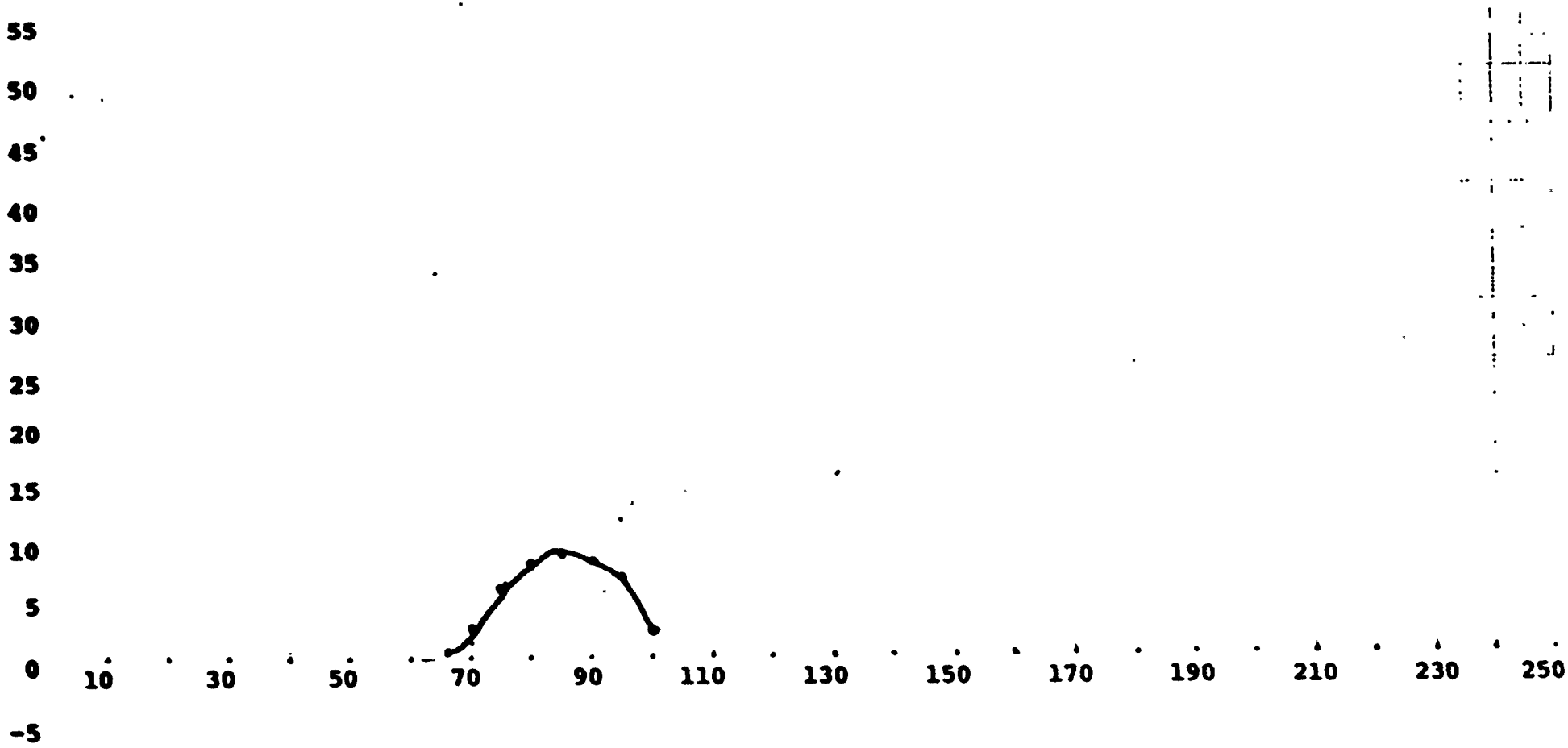

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 5


55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

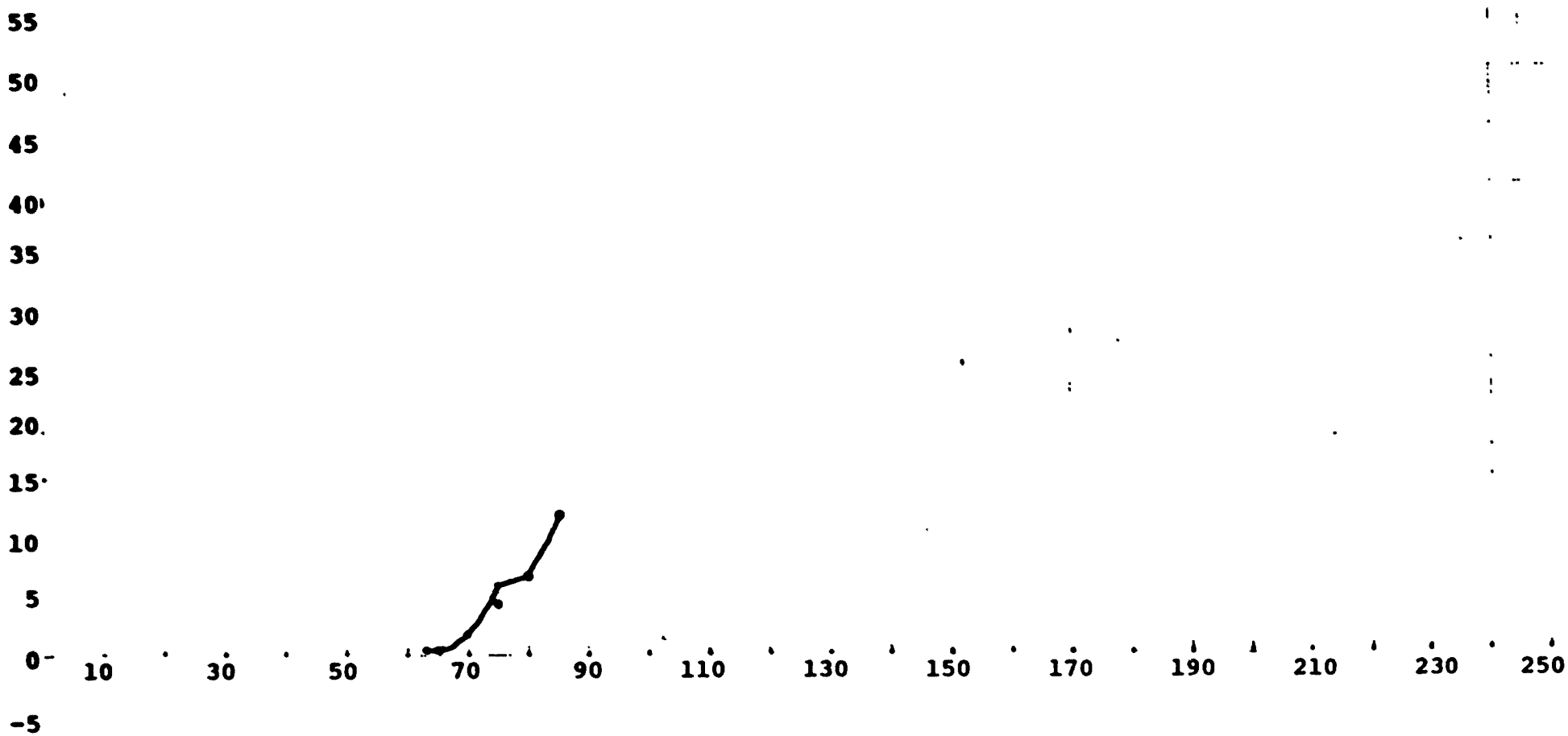

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 6

Figure 7

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figure 8

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 9

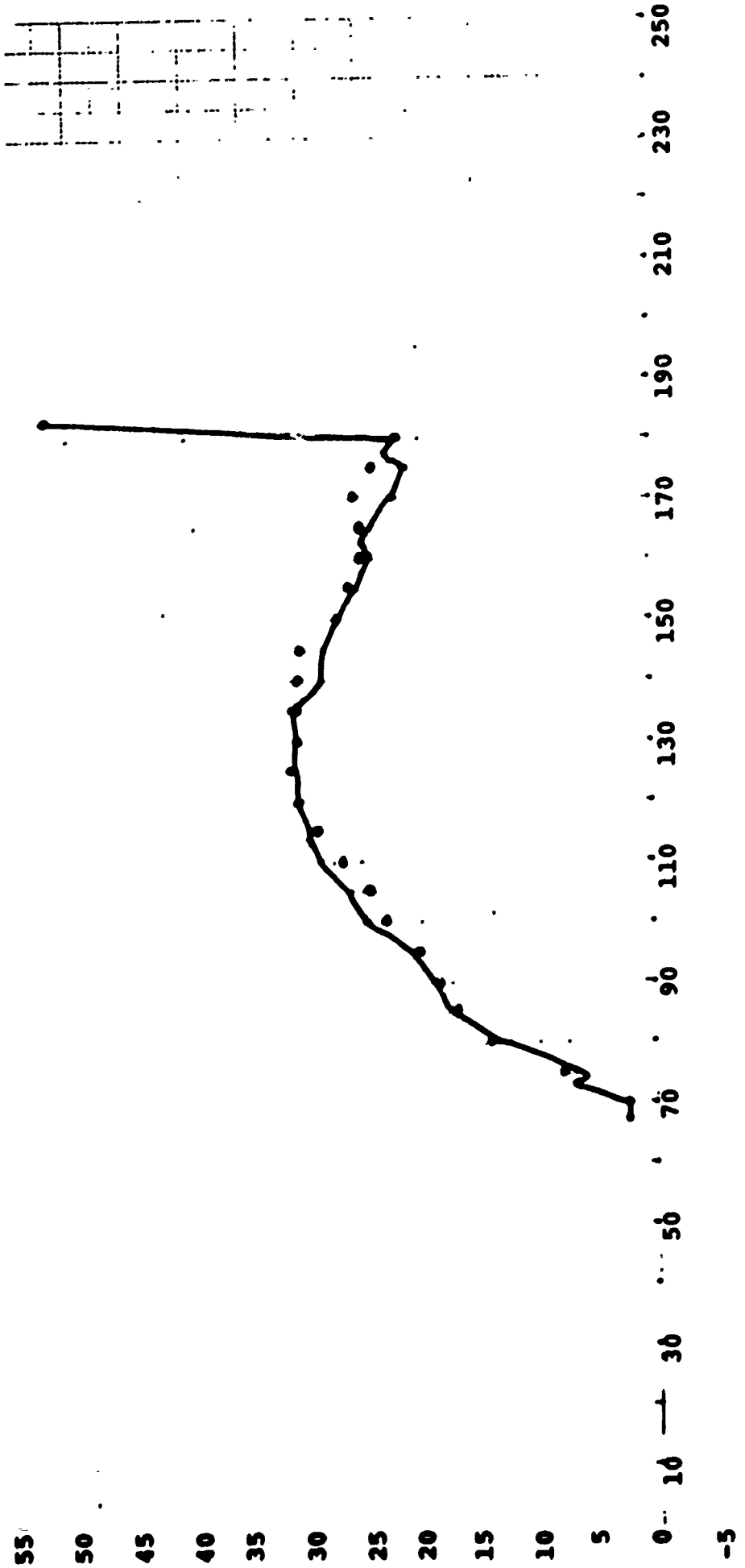
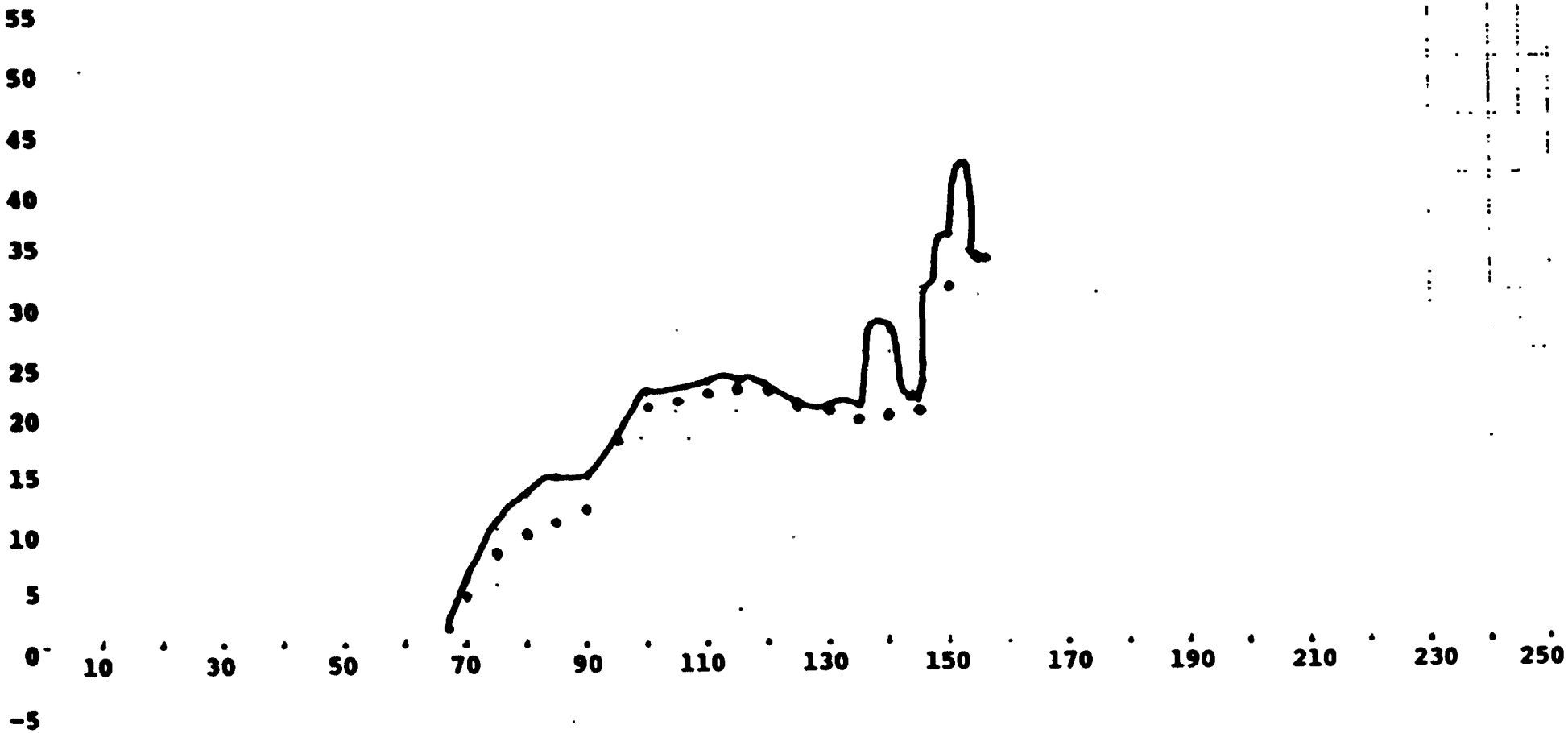

250
230
210
190
170
150
130
110
90
70
50
30
10
0
-5
55
50
45
40
35
30
25
20
15
10
5

Figure 10


55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 11


55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

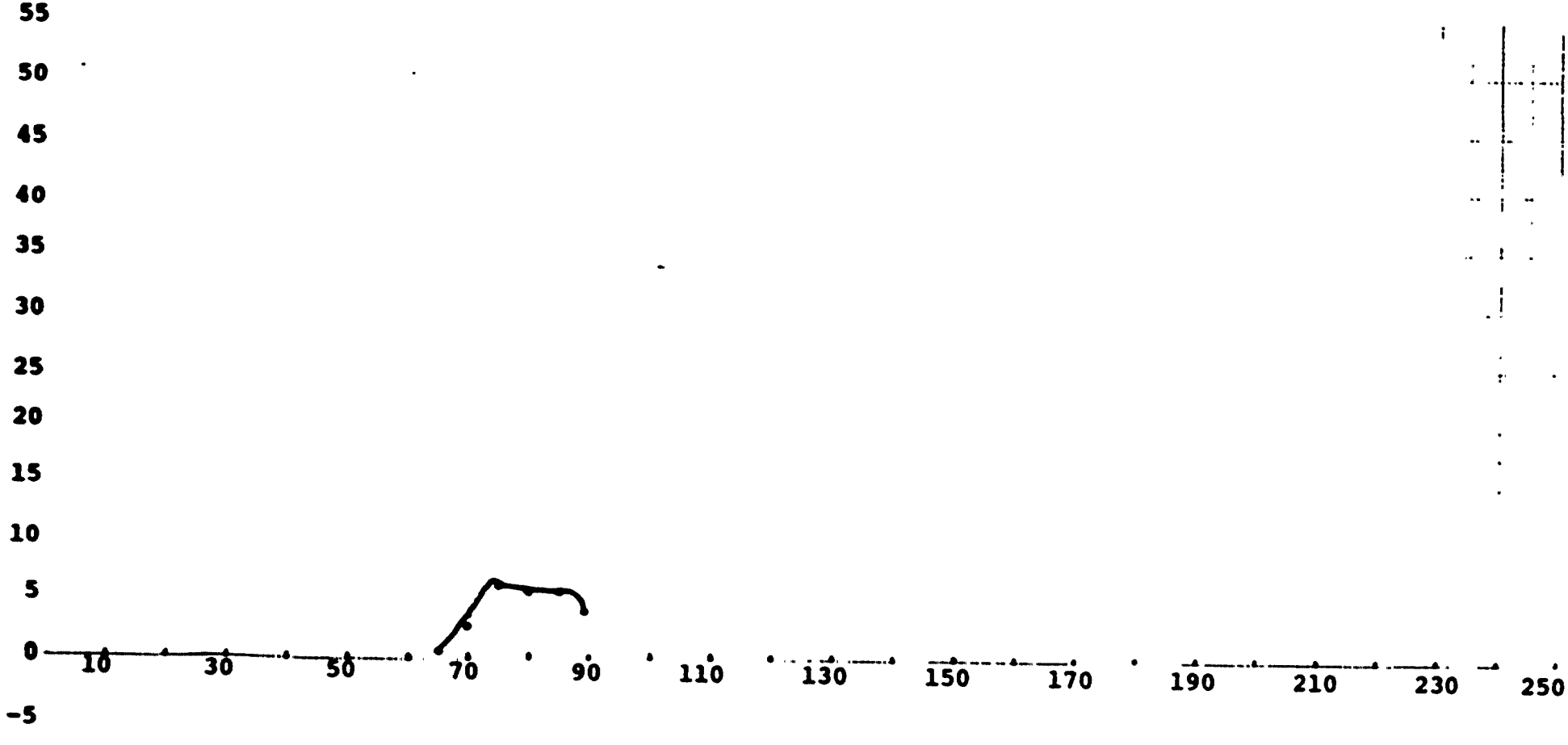

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 12

Figures 13-24

Border edge strength feature values of Samples Nos. 1-12 for all connected components using minimum edge operator.

Figure 13

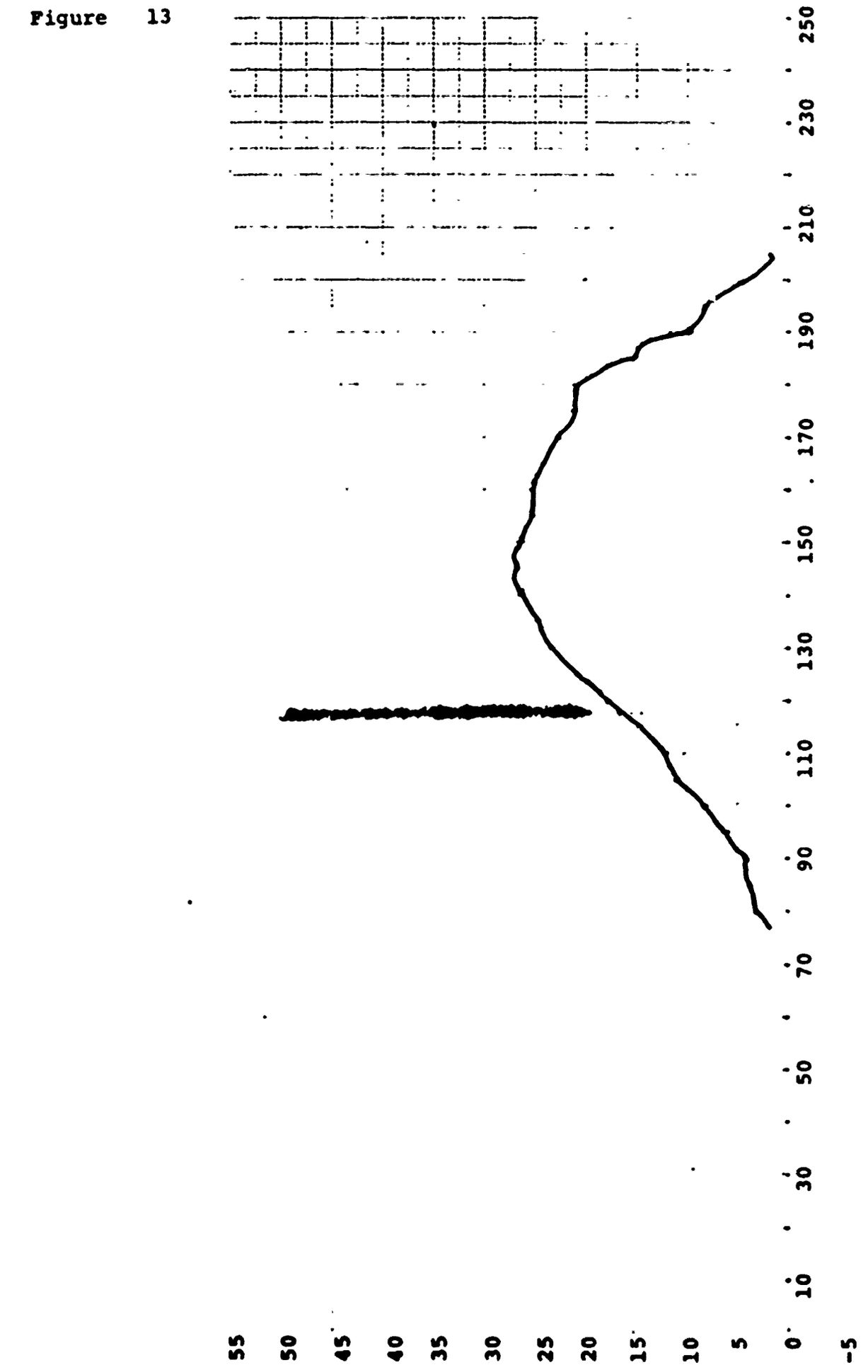
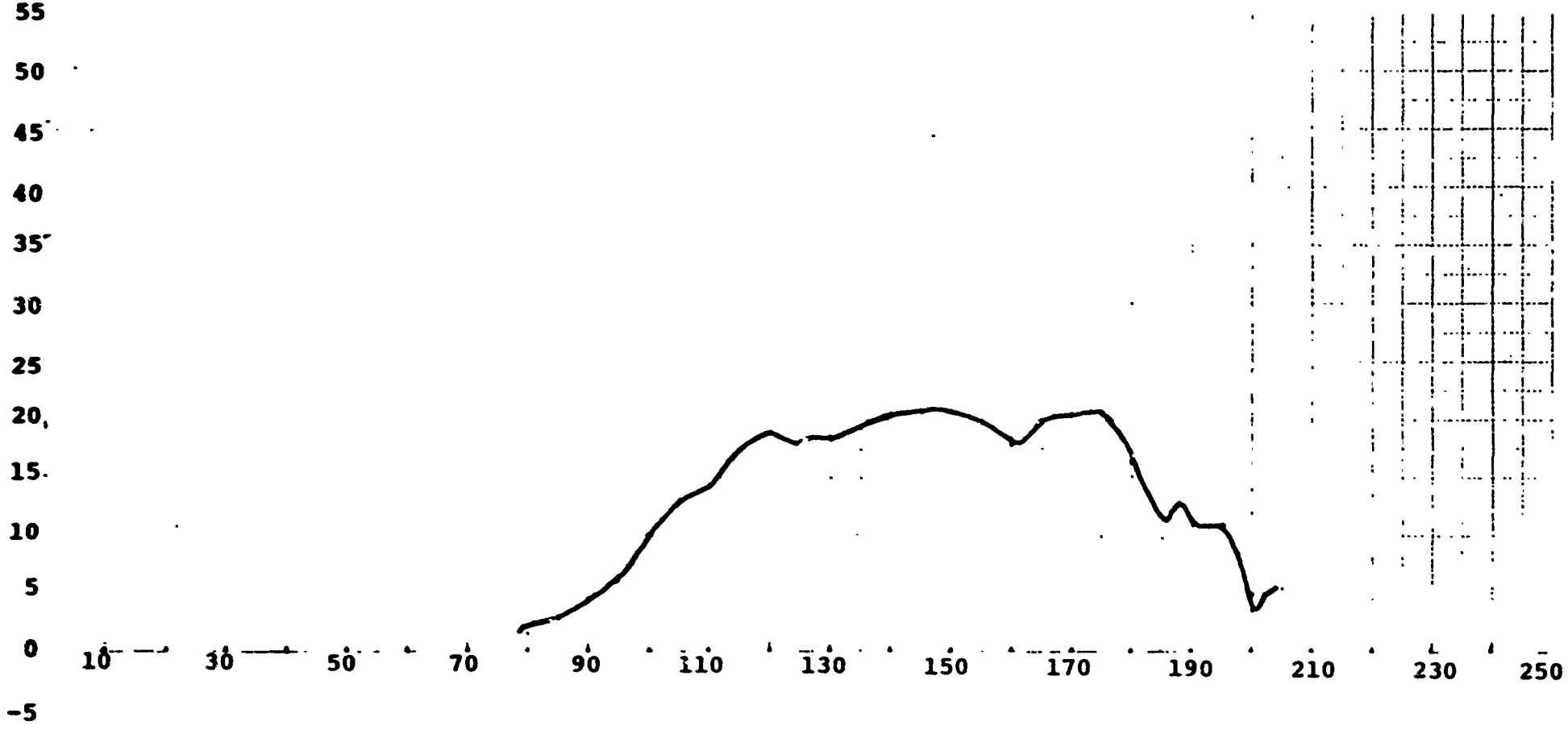

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figure 14


55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

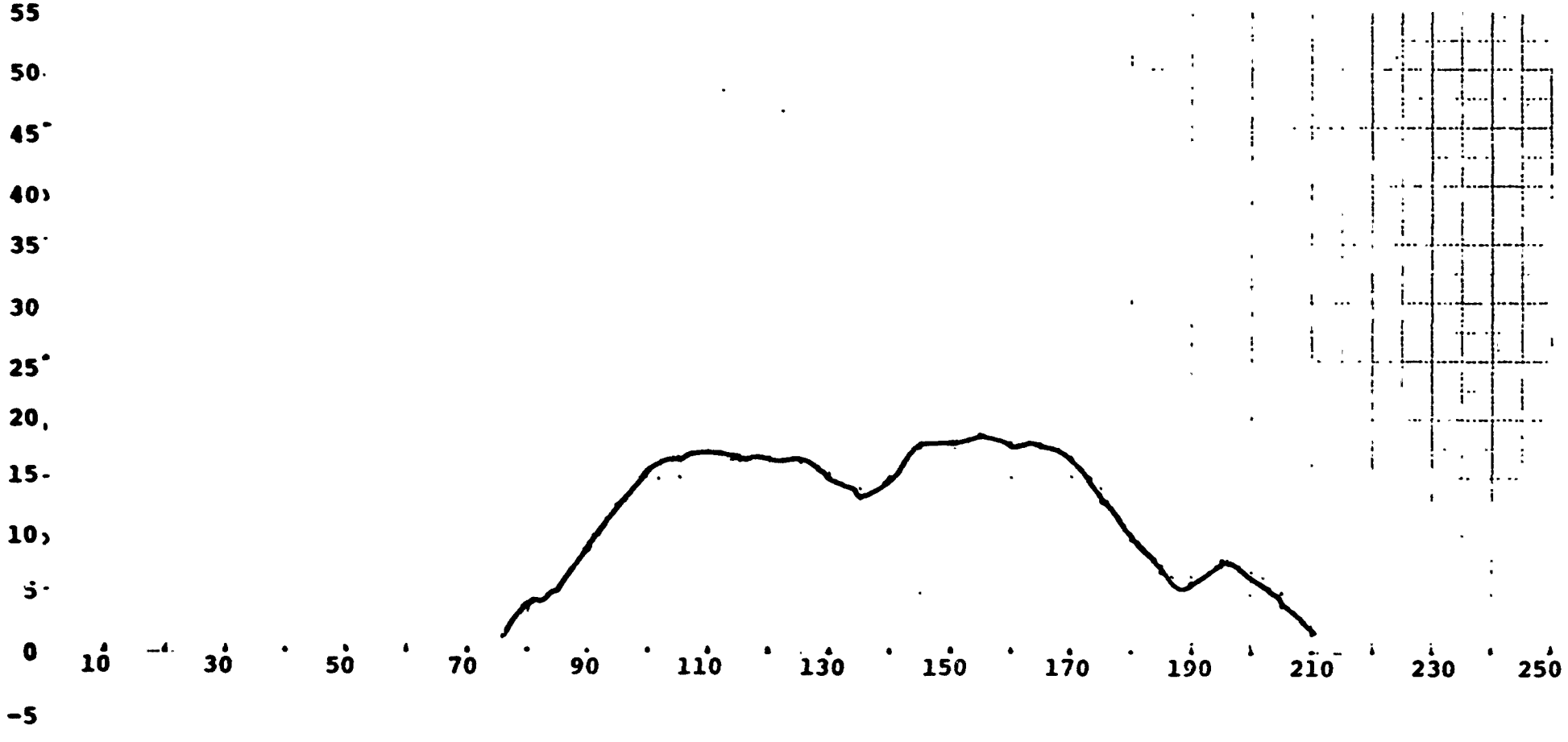

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 15

Figure 16

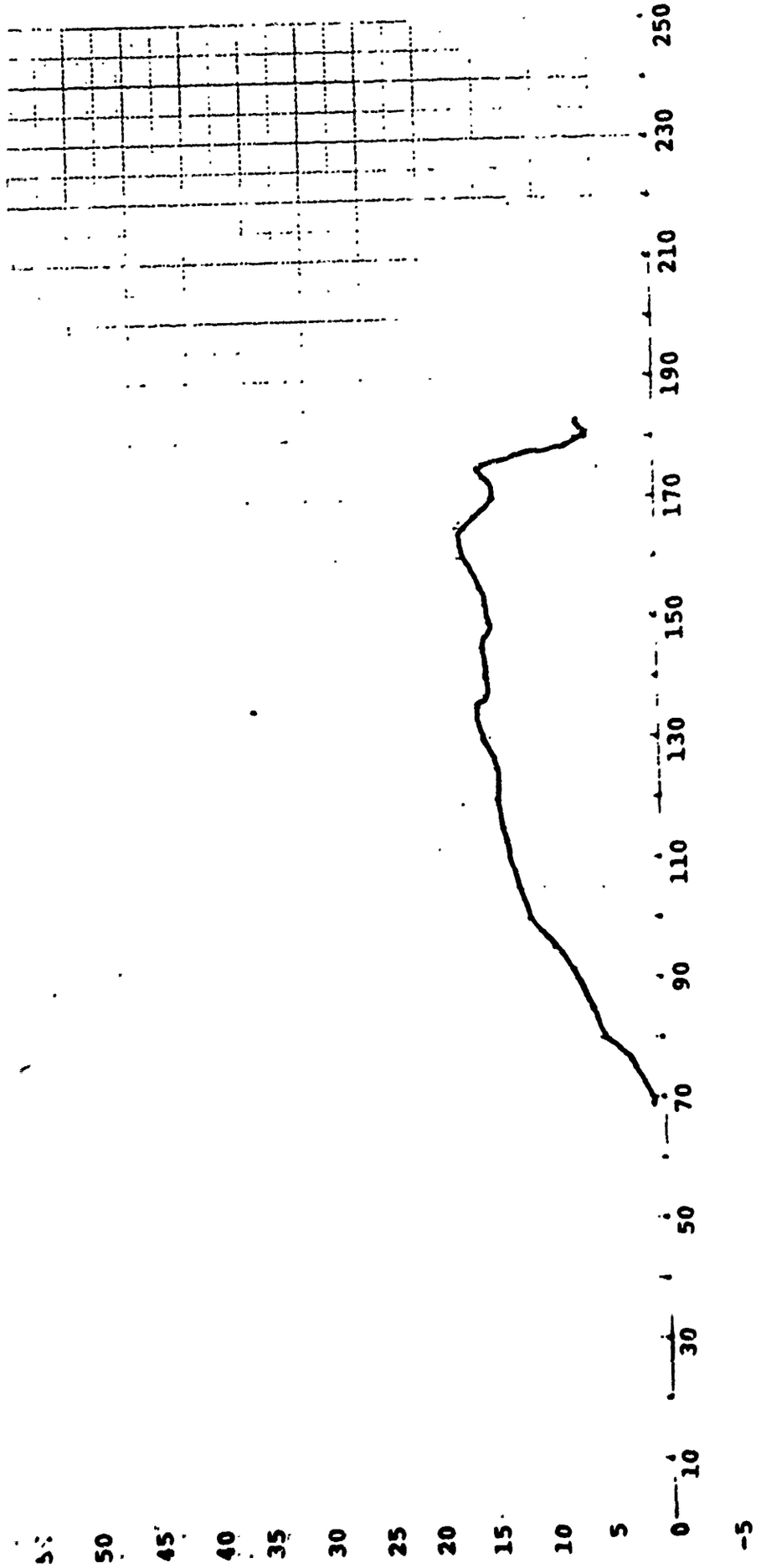

250
230
210
190
170
150
130
110
90
70
50
30
10
0
-5
5
50
45
40
35
30
25
20
15
10
5

Figure 17

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figure 18

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 19

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figure 20

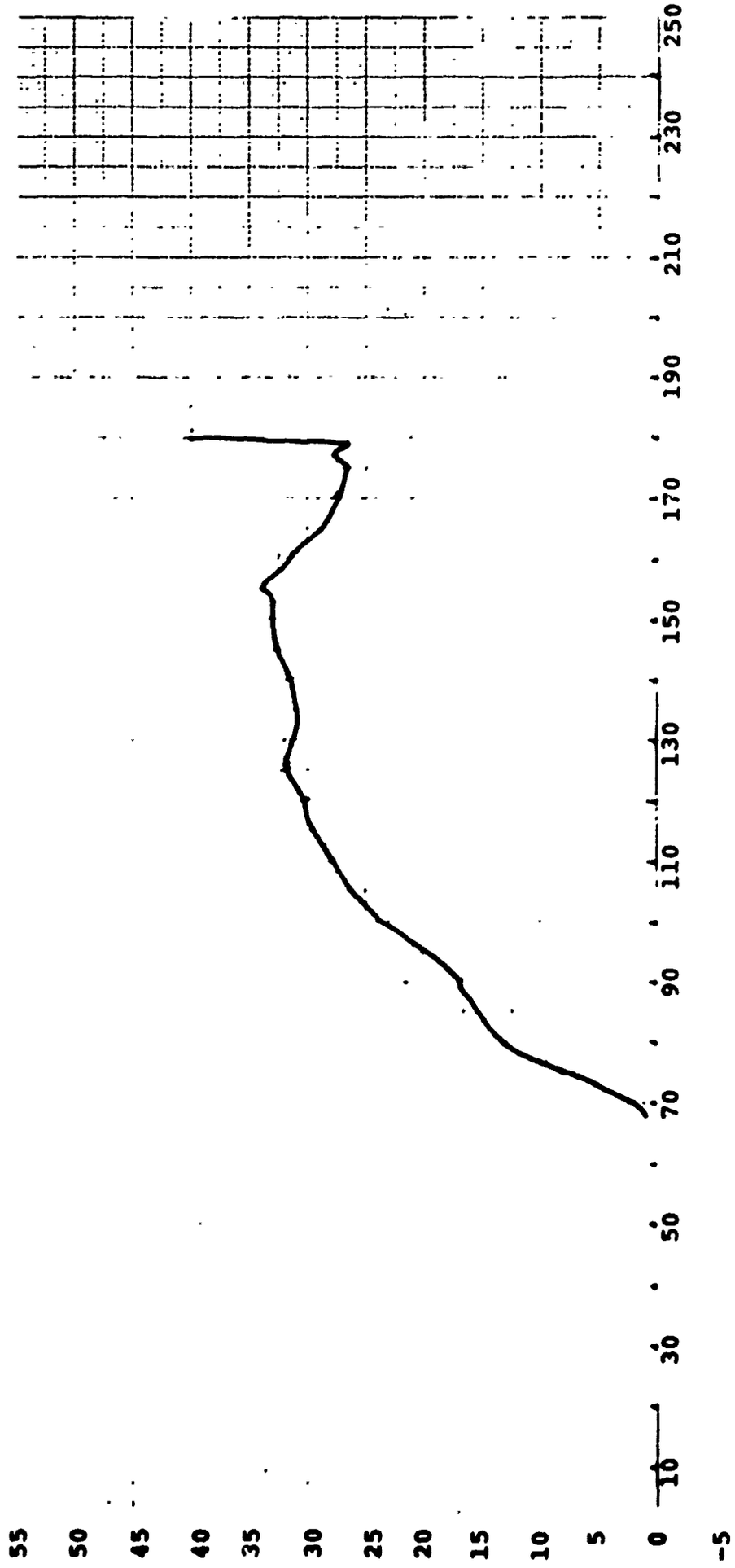

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figure 21

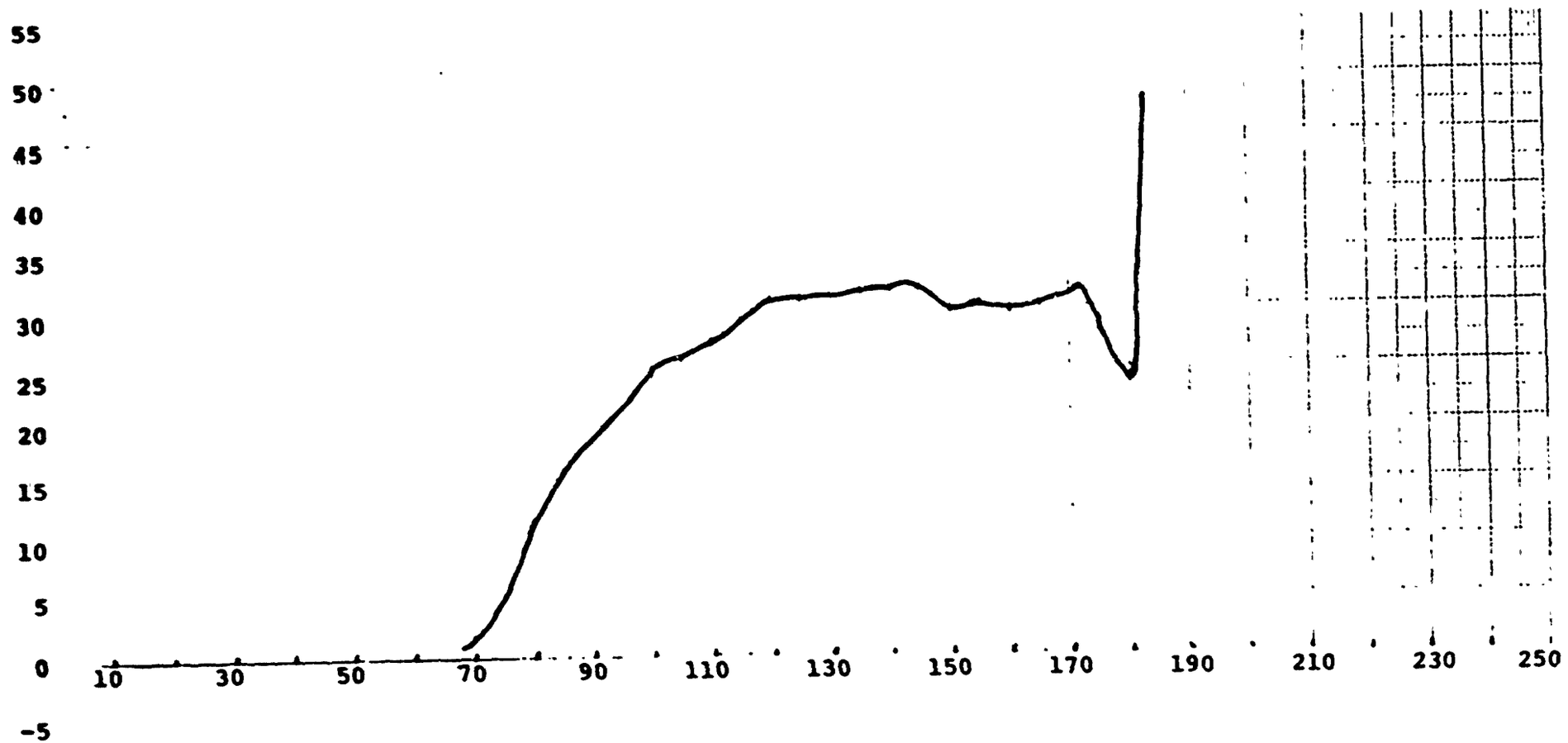

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 22

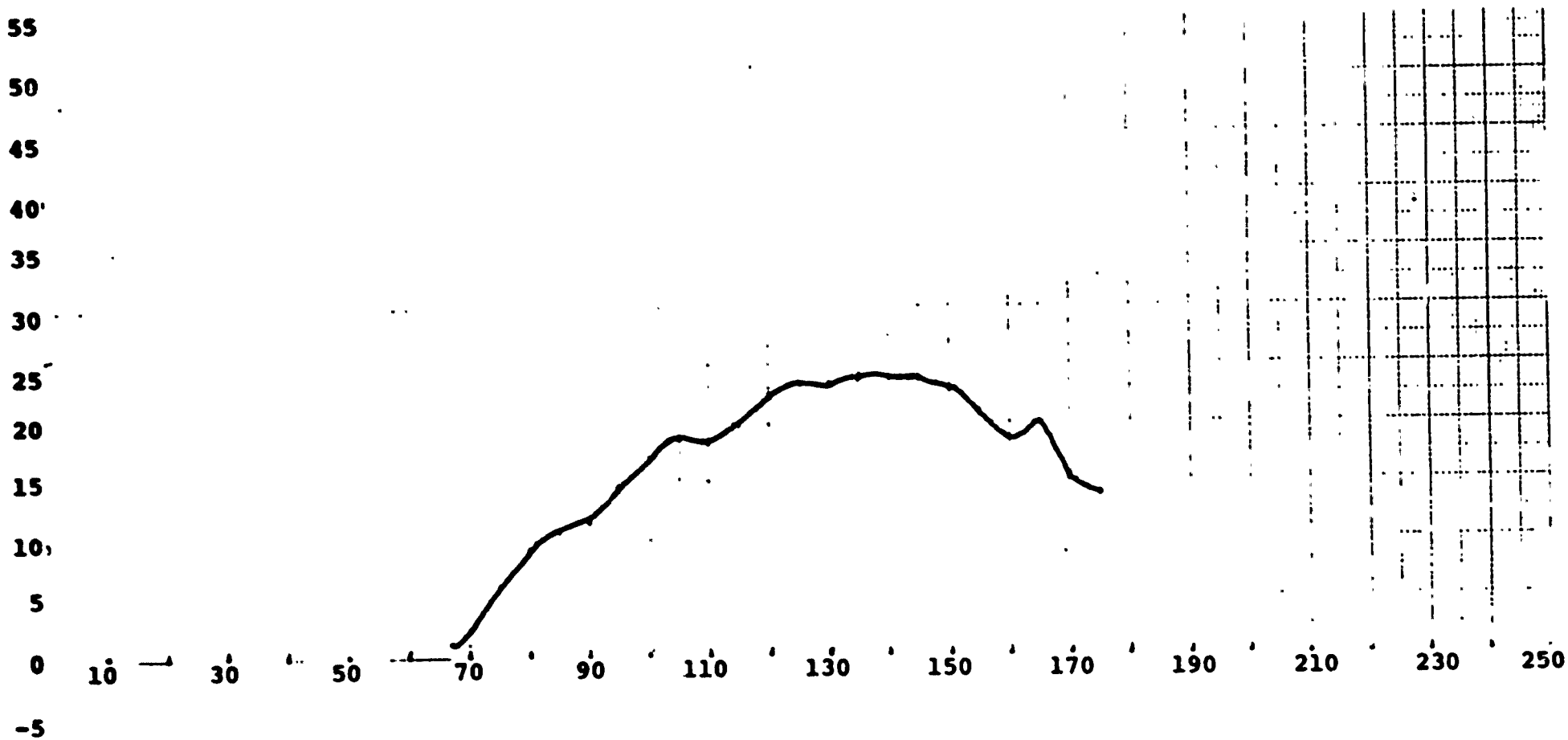

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 23

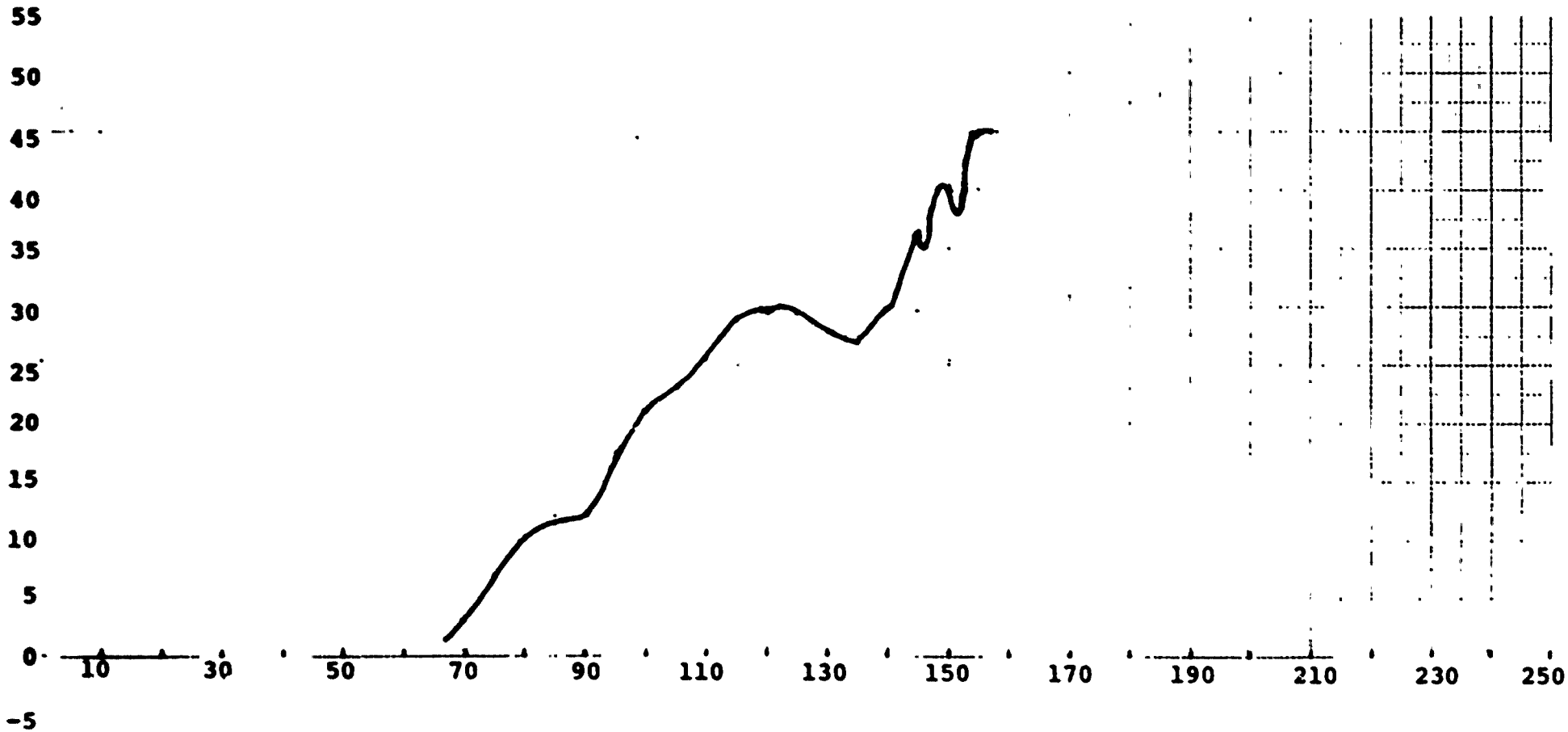

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 24

250
230
210
190
170
150
130
110
90
70
50
30
10
0
-5
55
50
45
40
35
30
25
20
15
10
5

Figures 25-36

Border edge strength feature values of Samples Nos. 1-12 for all connected components using average edge operator.

Figure 25

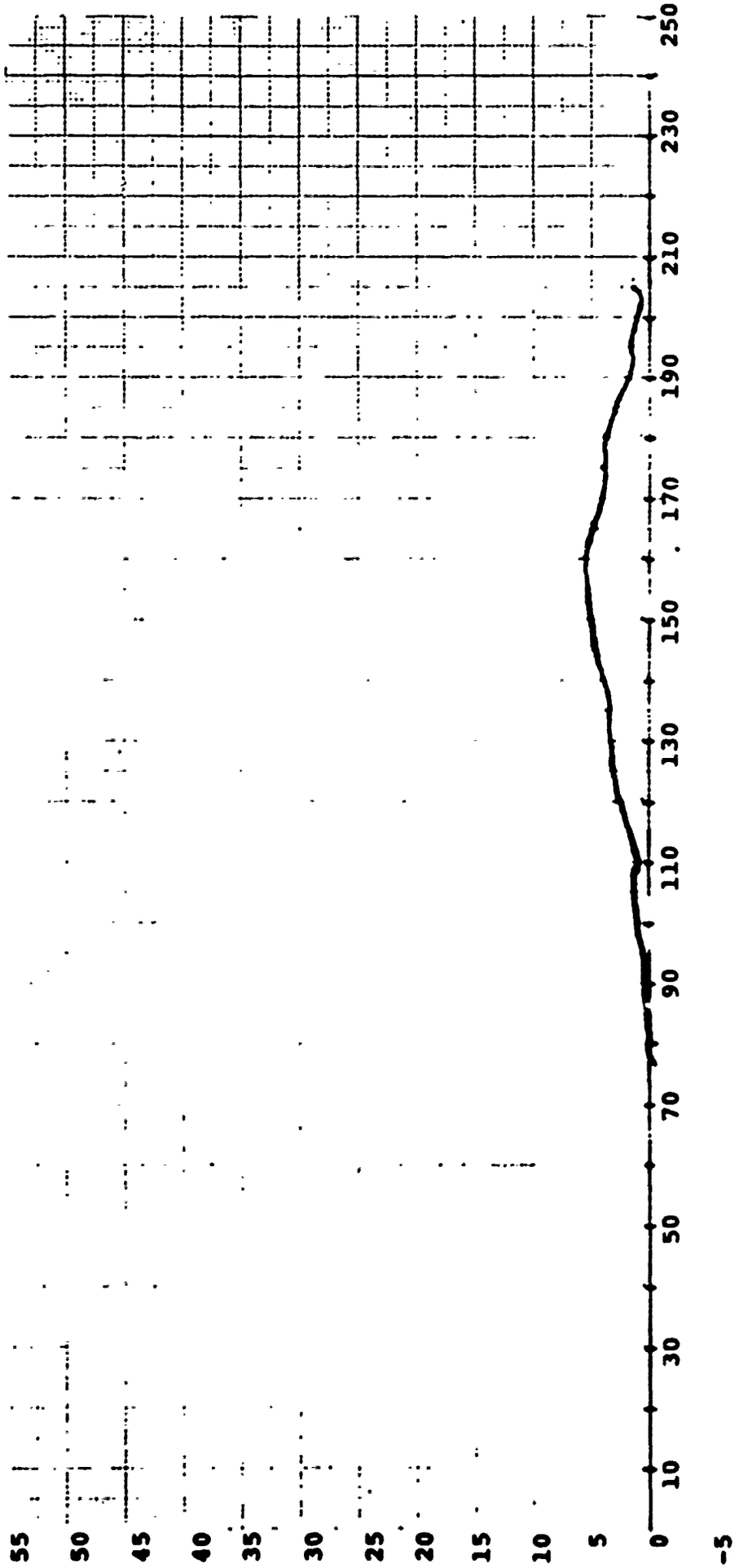

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figure 26

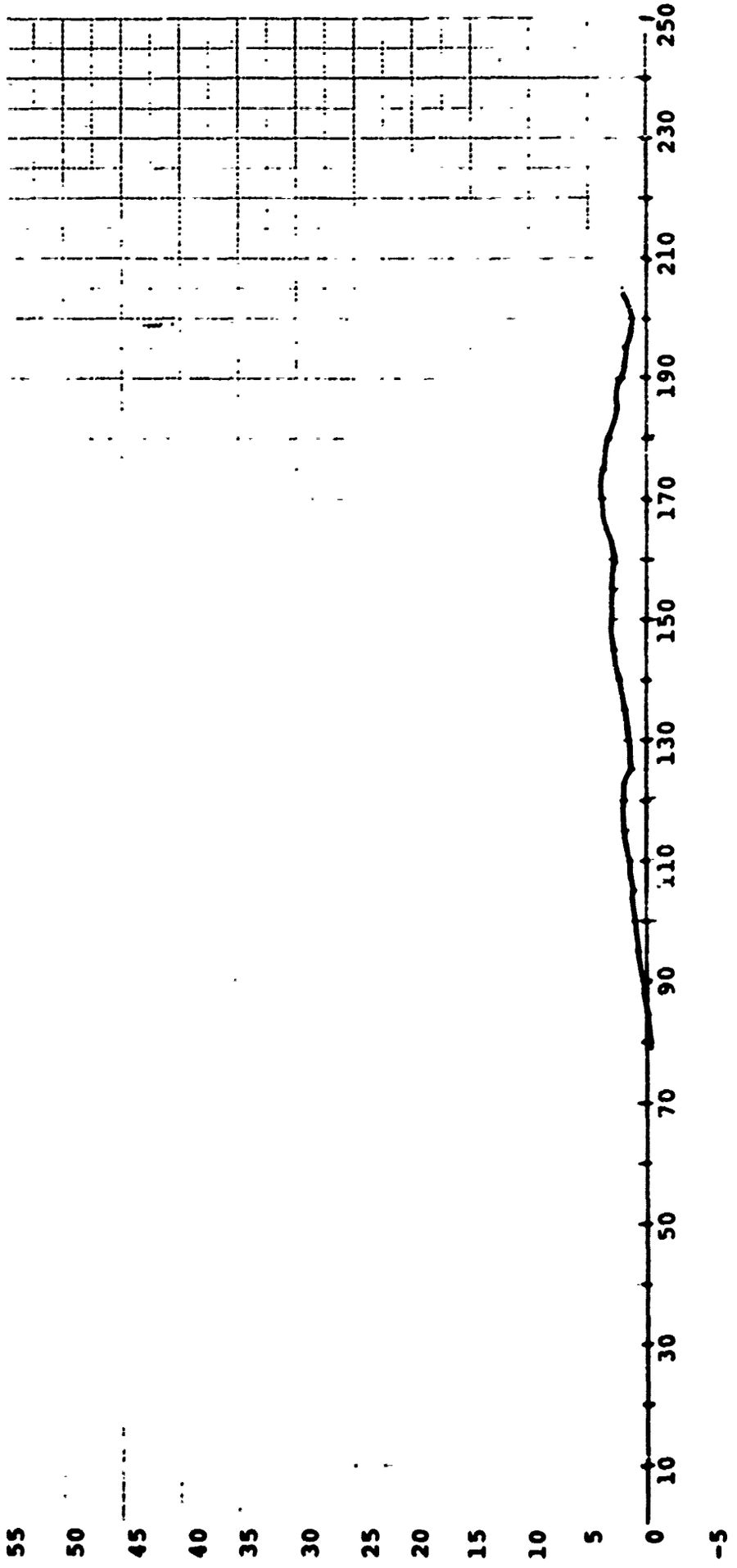

250
230
210
190
170
150
130
110
90
70
50
30
10
0
55
50
45
40
35
30
25
20
15
10
5
-5

Figure 27

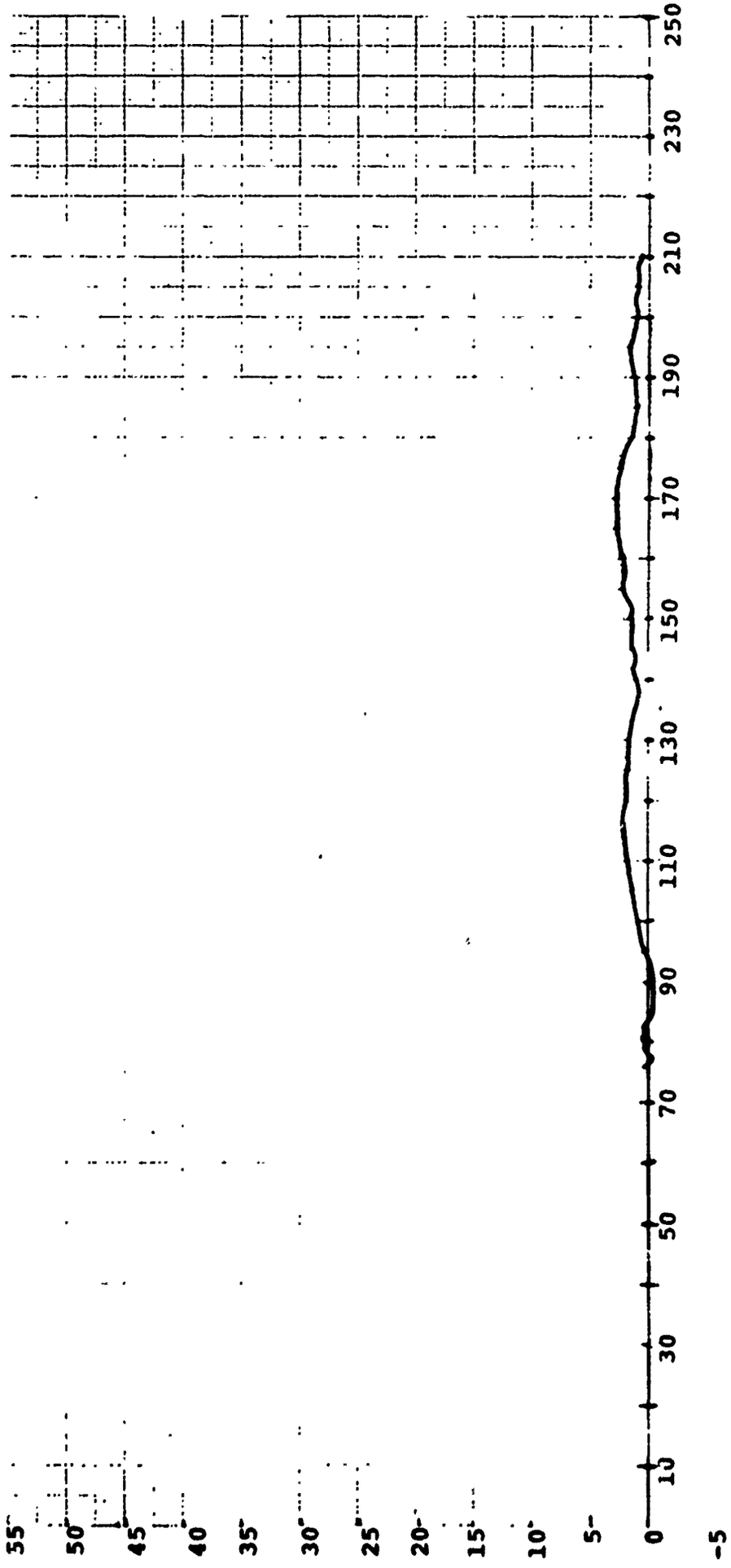

250
230
210
190
170
150
130
110
90
70
50
30
10
-5
0
5
10
15
20
25
30
35
40
45
50
55

Figure 28

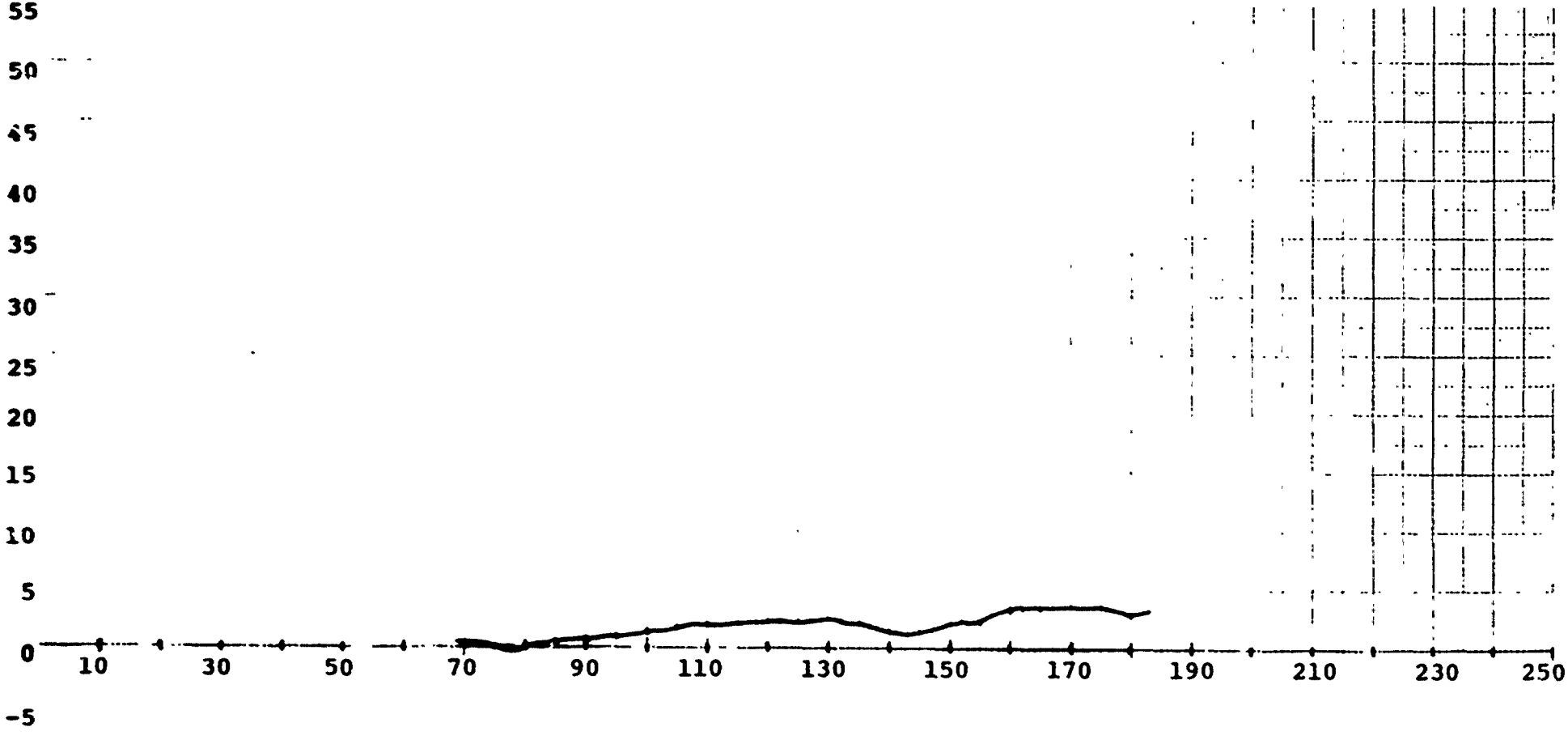

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 29

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figure 30

250
230
210
190
170
150
130
110
90
70
50
30
10
-5
0
5
10
15
20
25
30
35
40
45
50
55

Figure 5.31

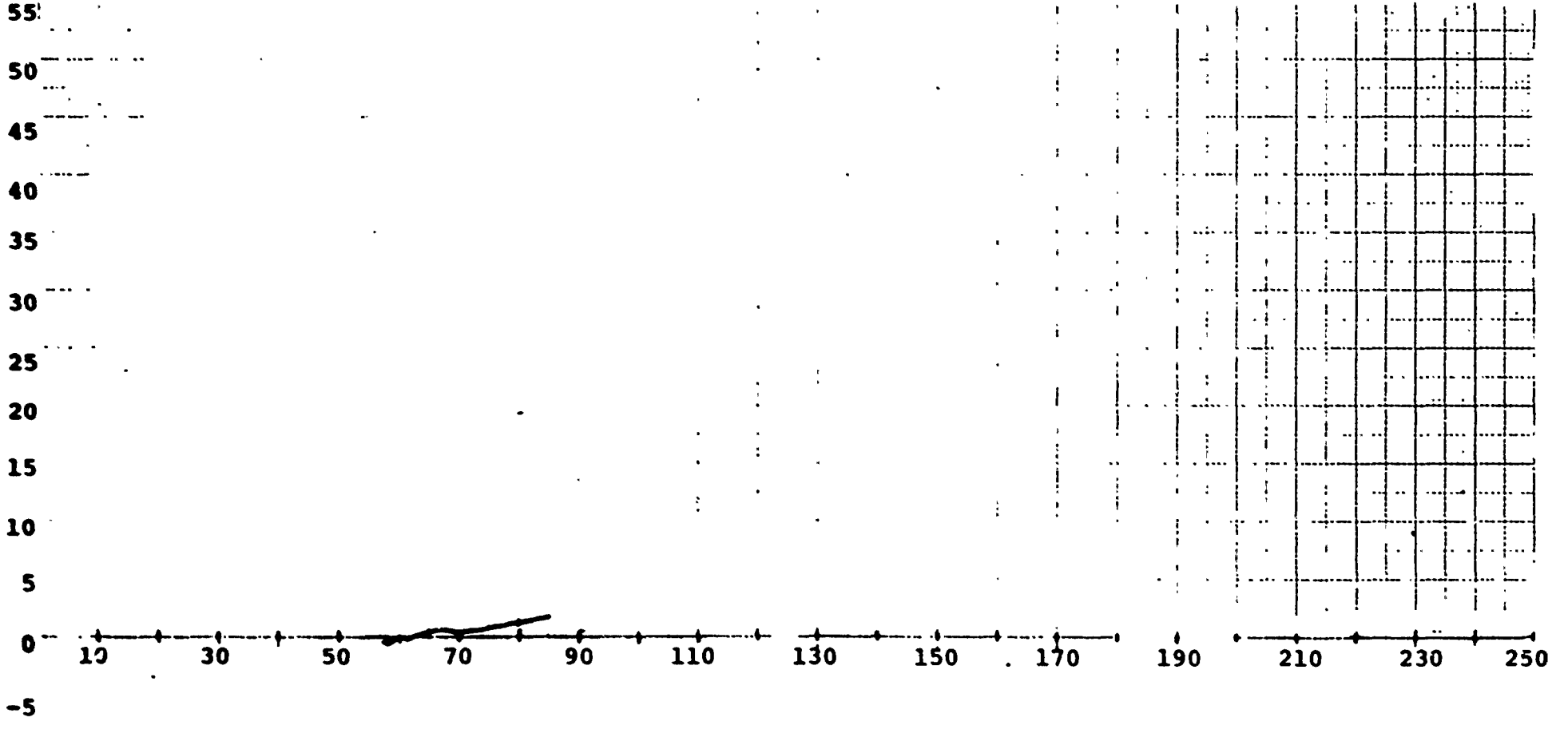

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 32

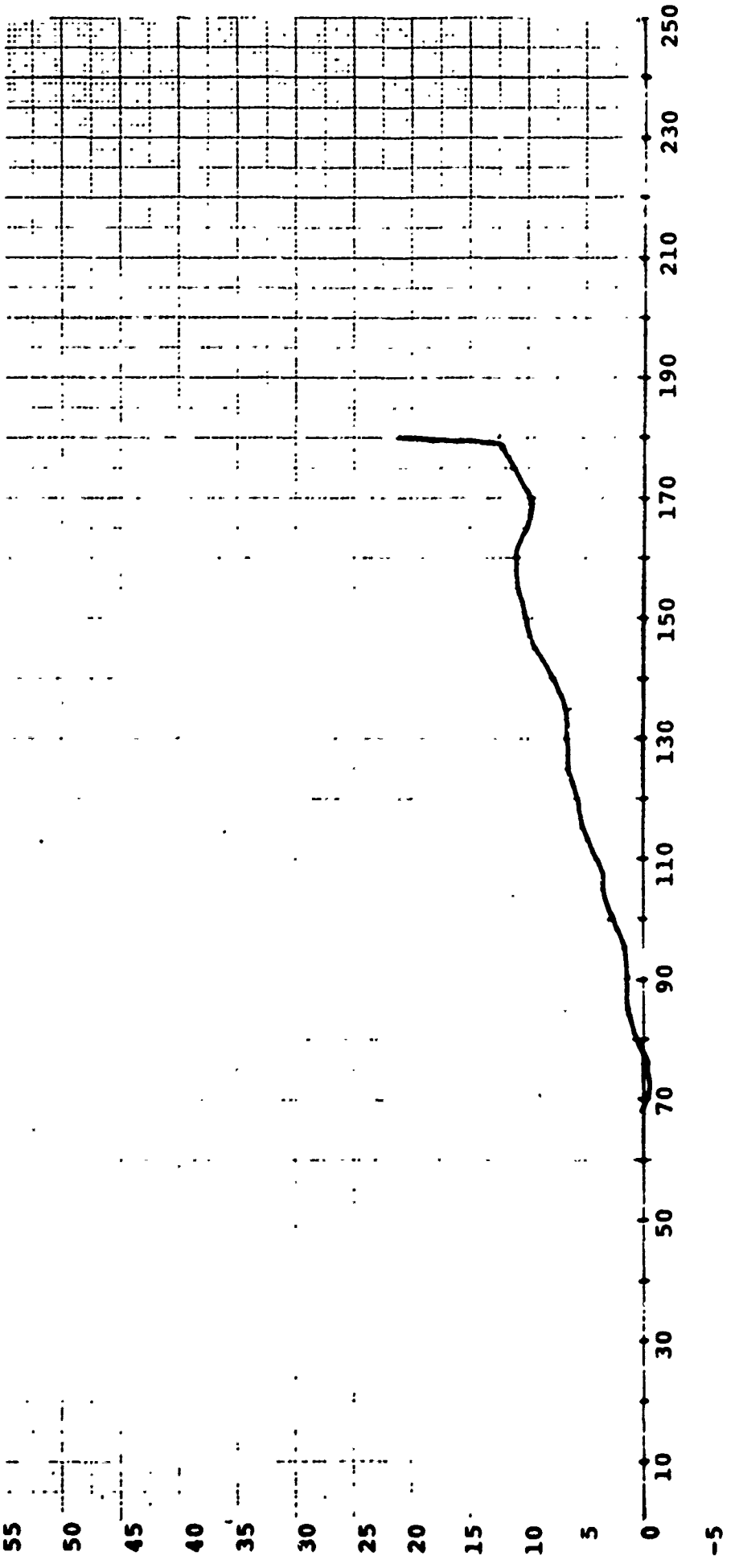

250
230
210
190
170
150
130
110
90
70
50
30
10
0
-5
55
50
45
40
35
30
25
20
15
10
5

Figure 33

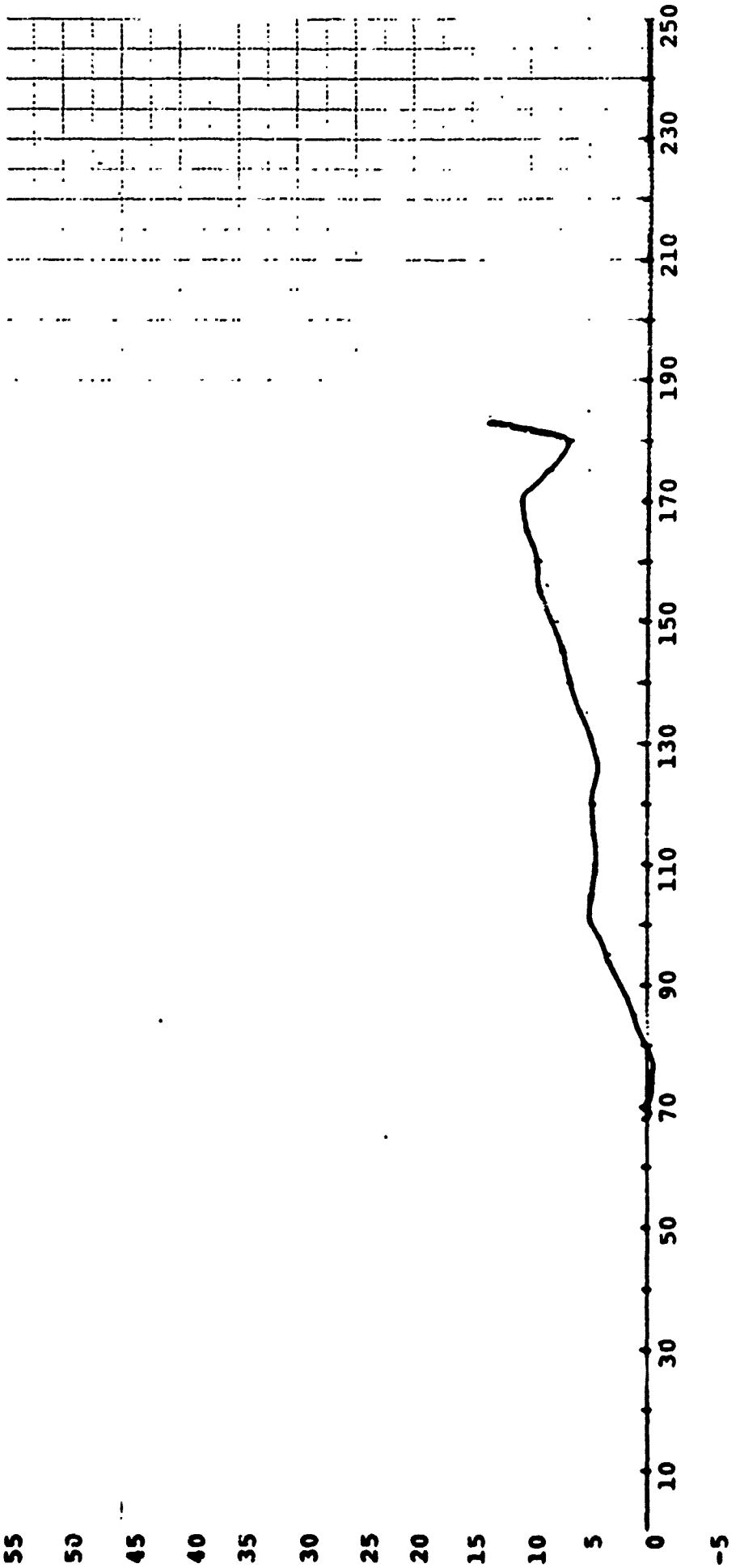

250
230
210
190
170
150
130
110
90
70
50
30
10
0
-5
55
50
45
40
35
30
25
20
15
10
5

Figure 34

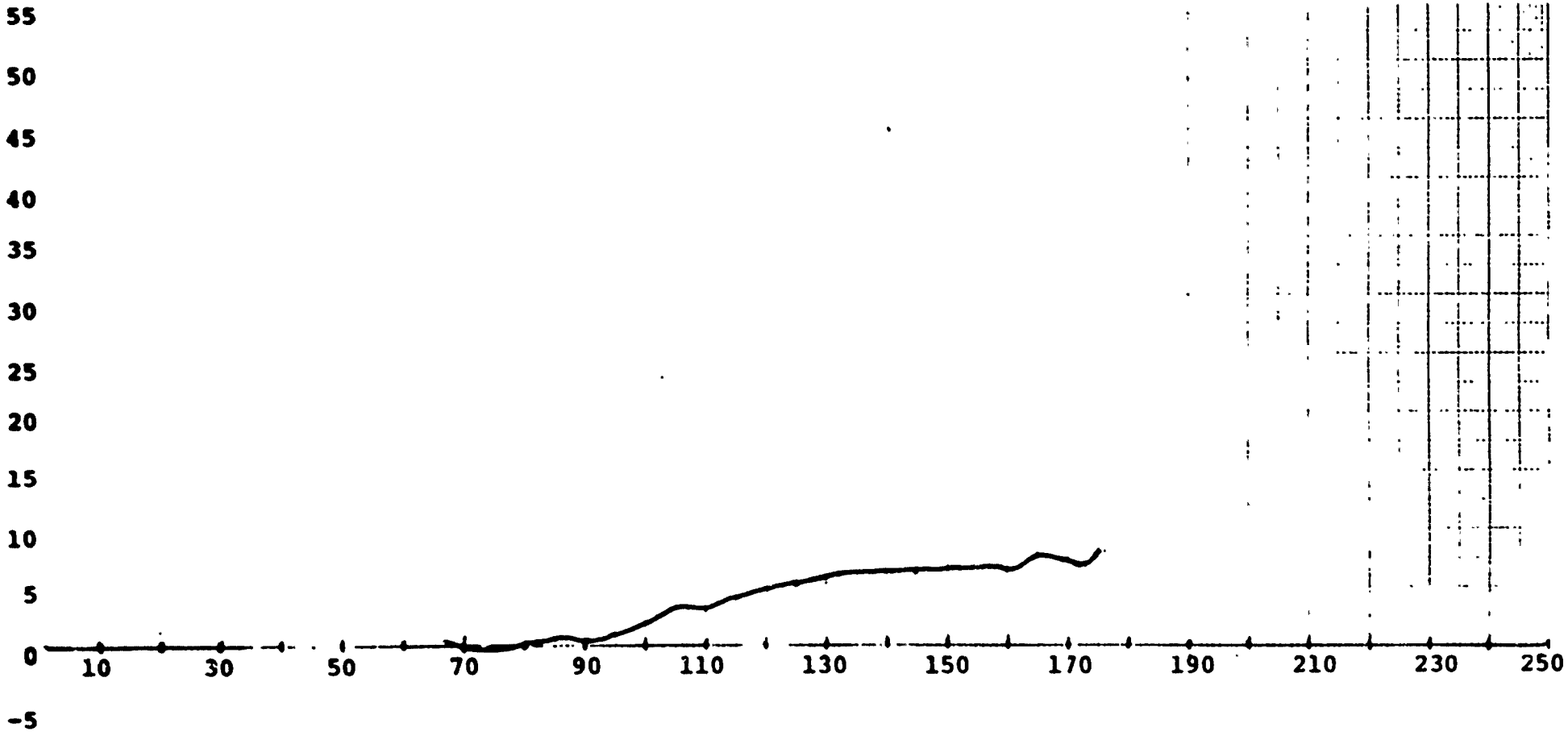

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 35

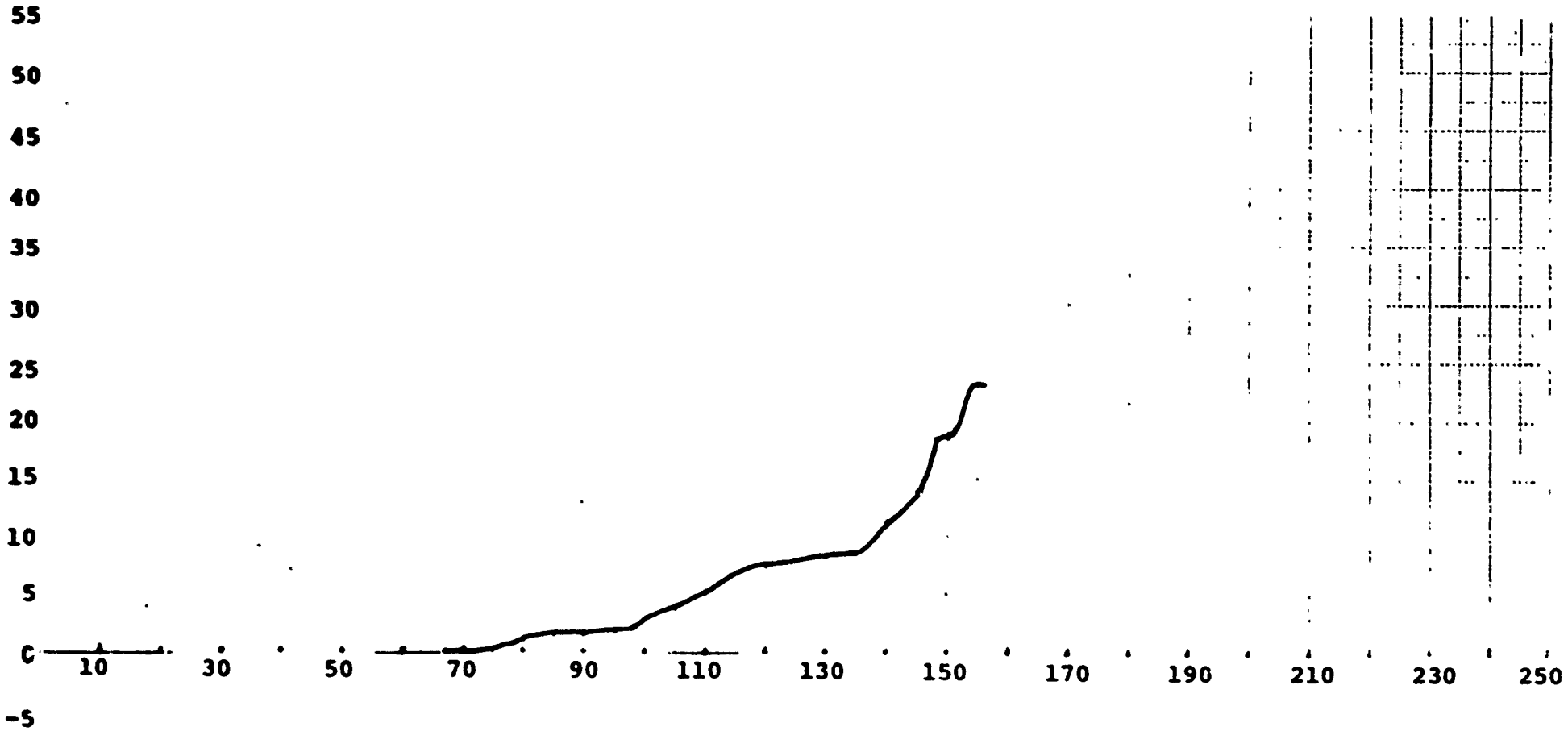

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 36

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figures 37-48

Border edge strength feature values of Samples Nos. 1-12 for all connected components using directional edge operator.

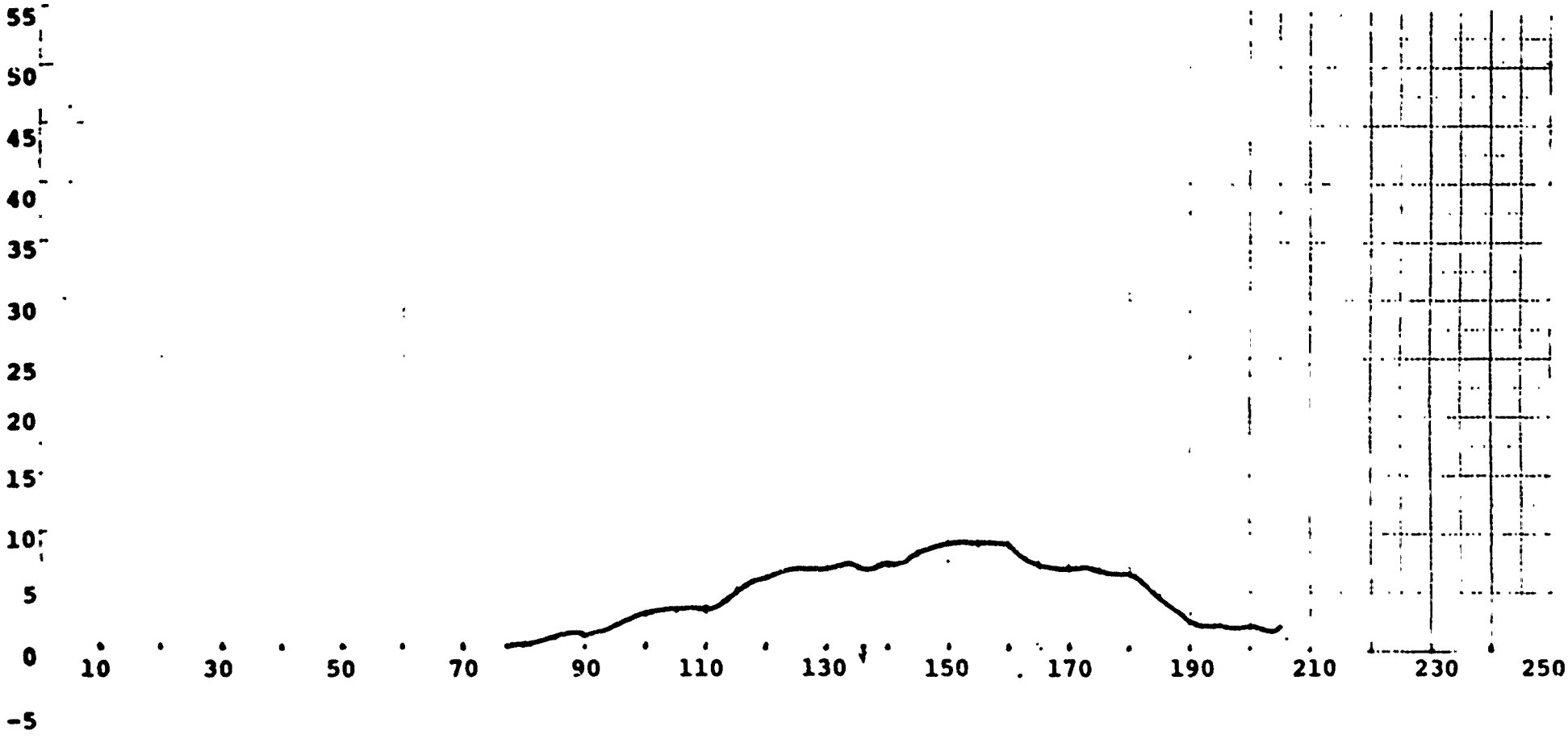

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 37

Figure 38

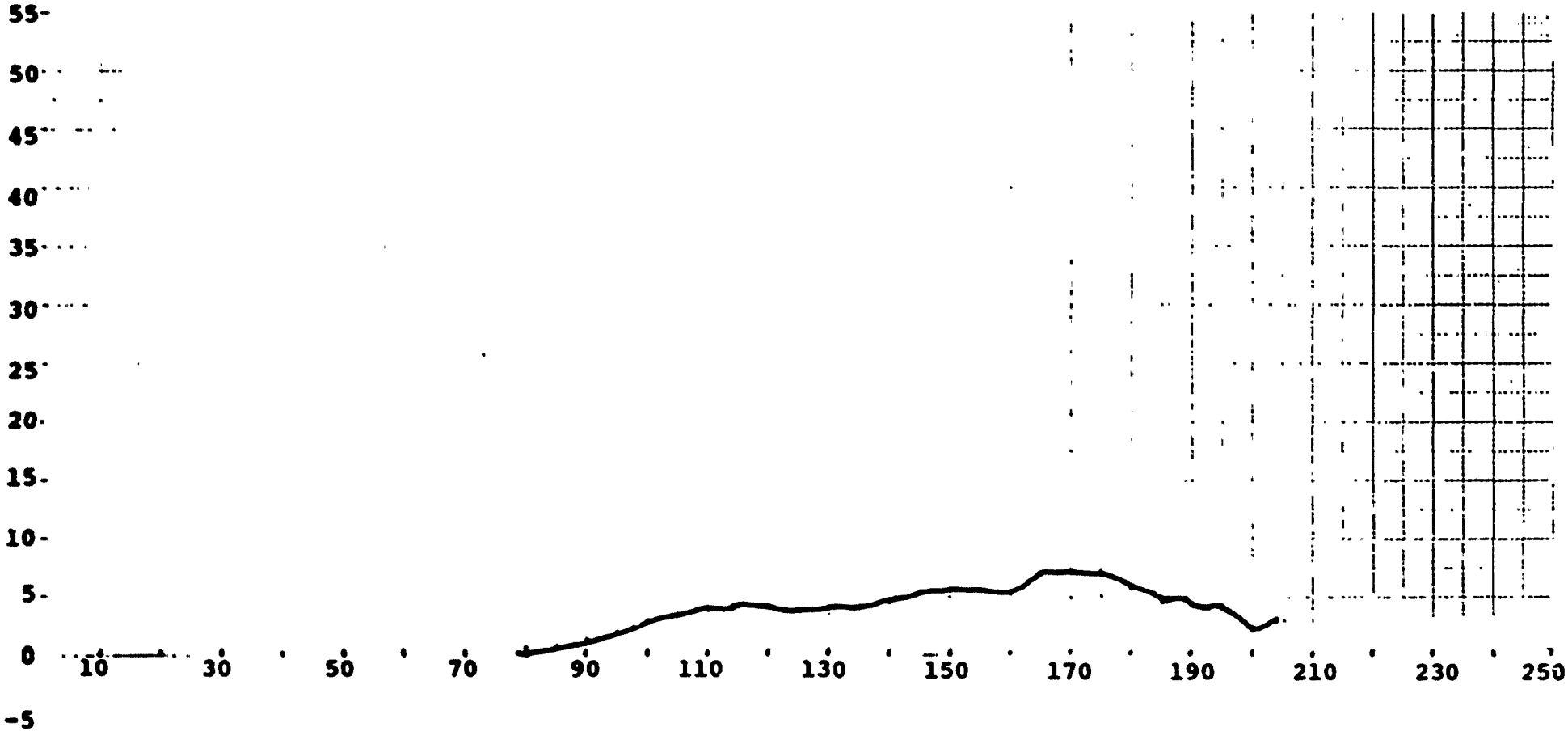

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 39

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figure 40

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 41

250
230
210
190
170
150
130
110
90
70
50
30
10
-5
0
5
10
15
20
25
30
35
40
45
50
55

Figure 42

250
230
210
190
170
150
130
110
90
70
50
30
10
-5
55
50
45
40
35
30
25
20
15
10
5
0

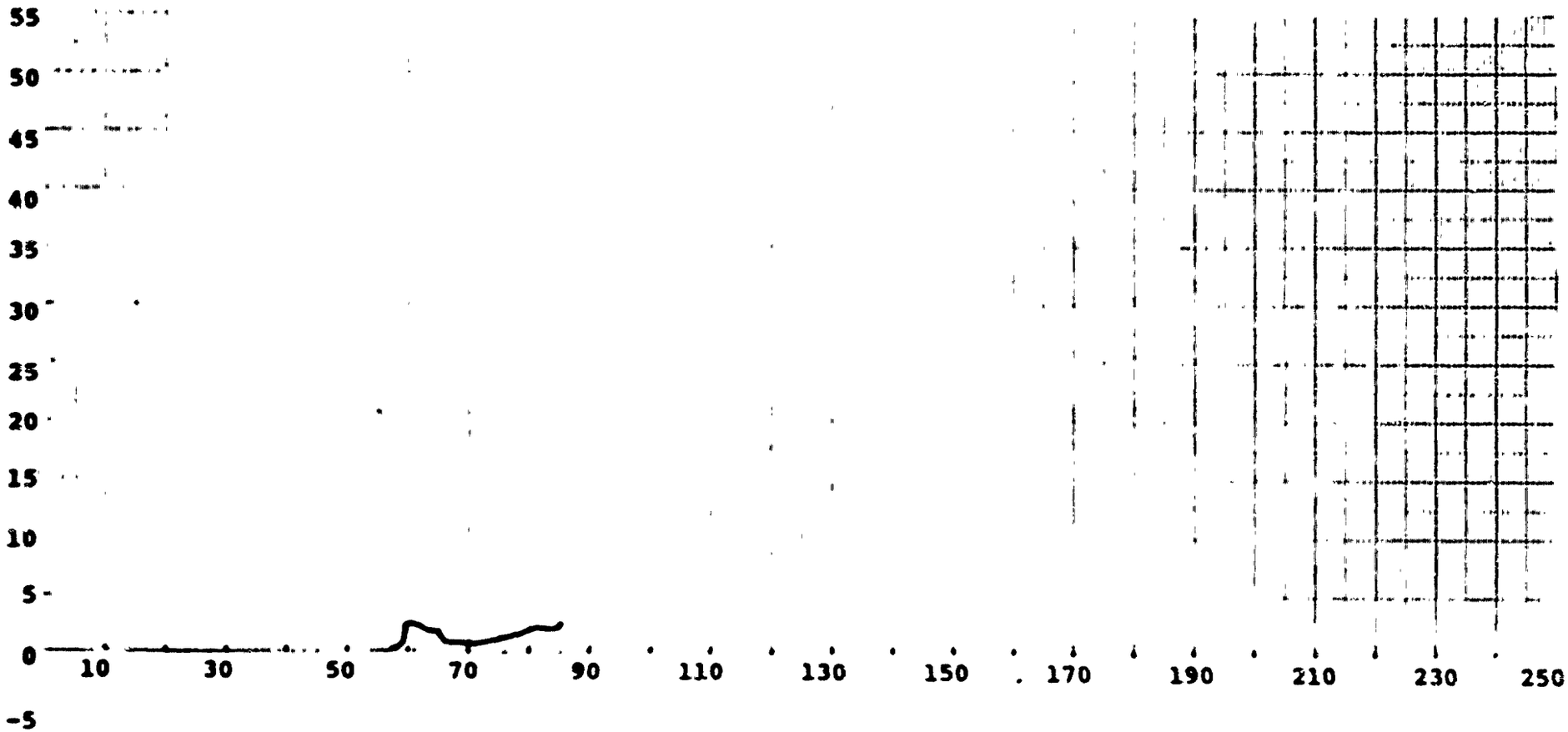

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 43

Figure 44

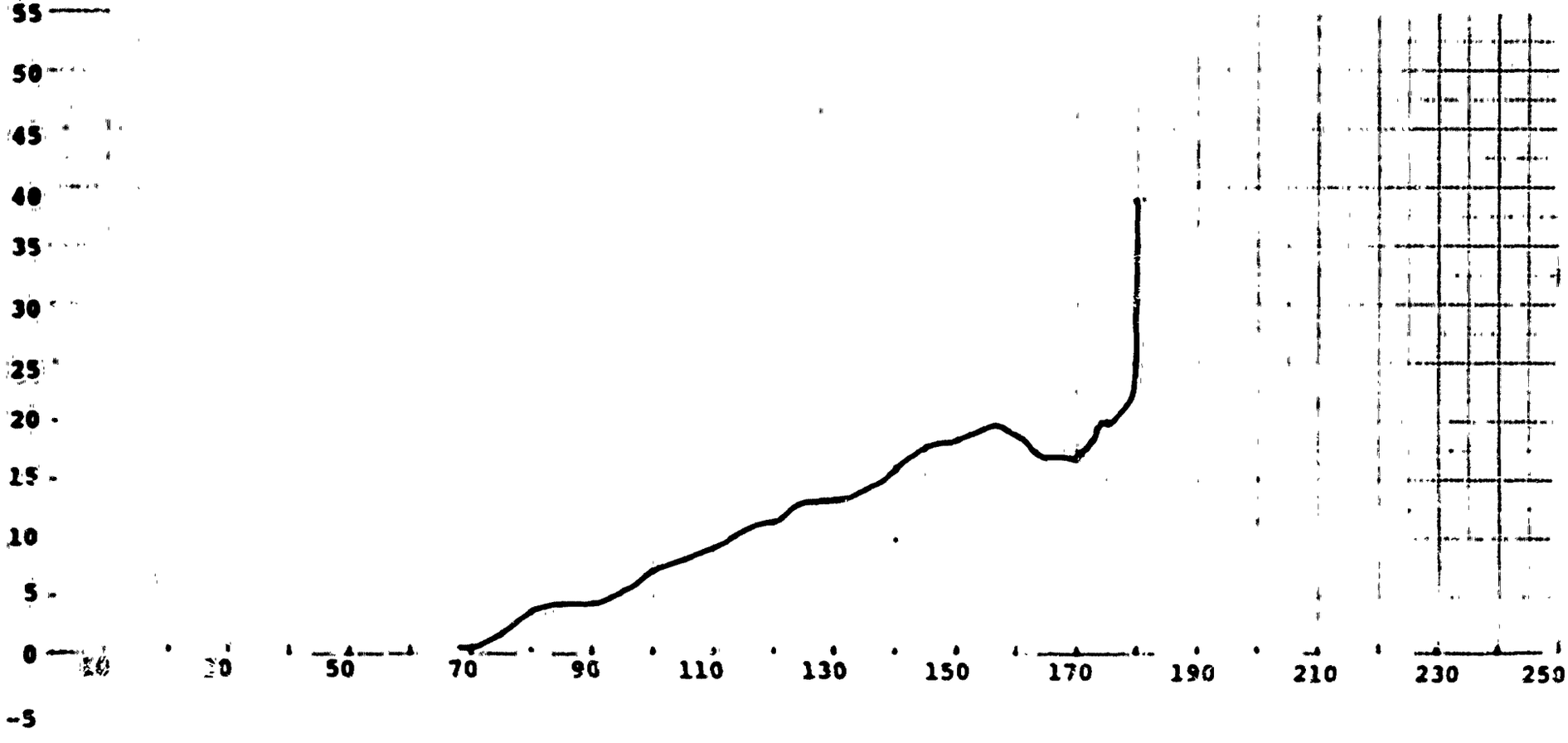

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 45

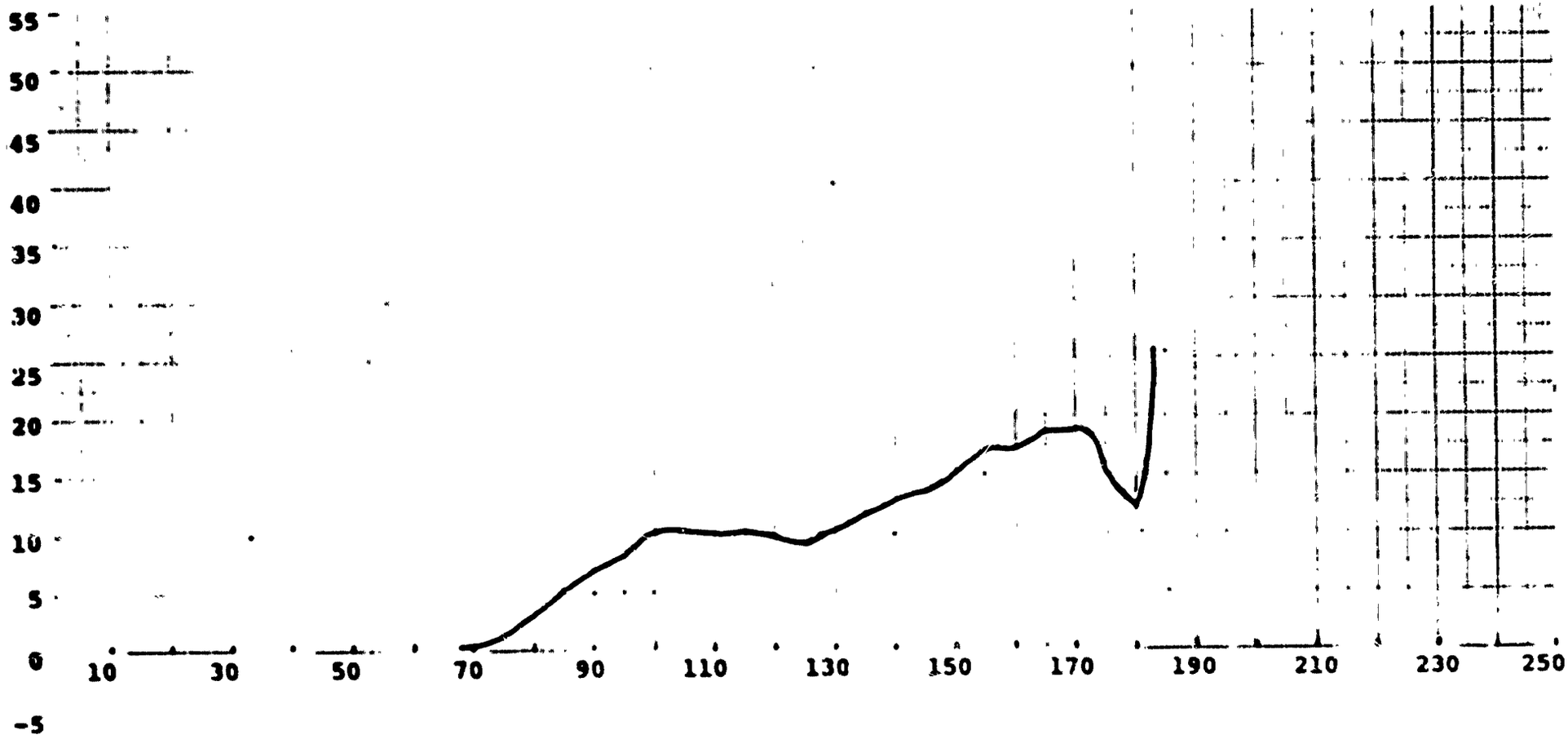

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 46

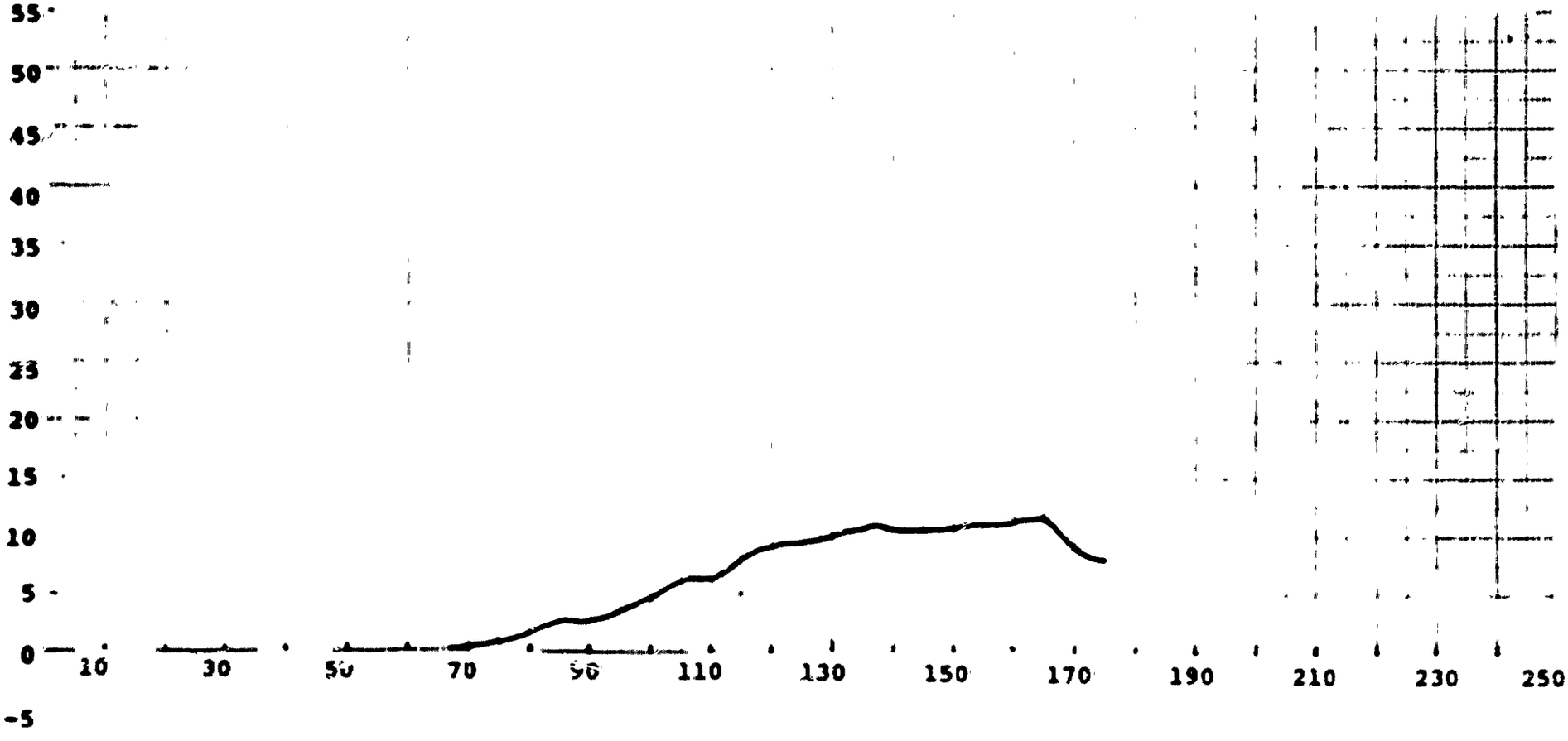

55
50
45
40
35
30
25
20
15
10
5
0
-5
10
30
50
70
90
110
130
150
170
190
210
230
250

Figure 47

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figure 48

250
230
210
190
170
150
130
110
90
70
50
30
10
55
50
45
40
35
30
25
20
15
10
5
0
-5

Figures 49-60

Infrared histograms for Samples Nos. 1-12, with coldest cluster threshold marked.

Figure 49

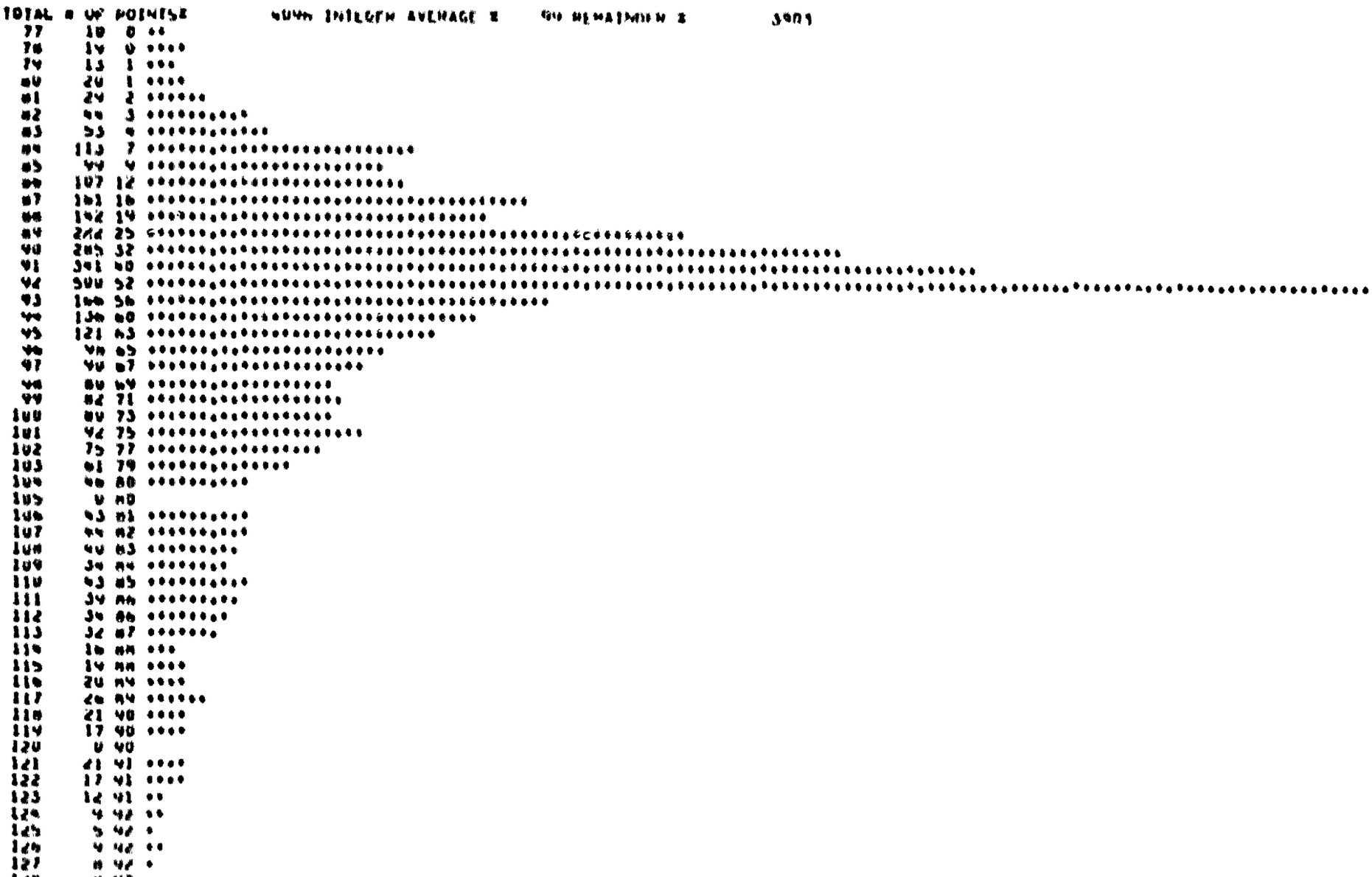


Figure 49 (cont'd)

Figure 49 (cont'd)

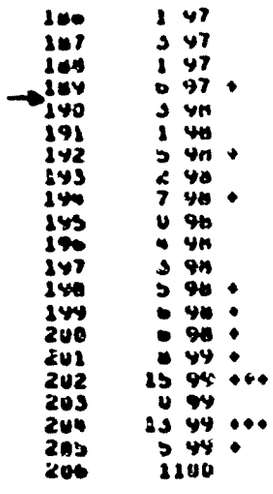


Figure 50


TOTAL # OF POINTS= 4096 INTEGER AVERAGE = 104 REMAINDER = 1020
79 4 0 **
80 14 0 ***
81 28 1 *******
82 53 2 *************
83 74 4 ******************
84 121 7 ******************************
85 121 10 ******************************
86 135 13 **********************************
87 133 16 **********************************
88 164 20 ***
89 144 24 ************************************
90 127 27 ********************************
91 158 31 **
92 471 42 **
93 175 47 **
94 137 58 **********************************
95 124 53 ********************************
96 120 56 ******************************
97 99 58 *************************
98 96 61 ************************
99 74 62 ******************
100 73 64 ******************
101 57 66 **************
102 50 67 ************
103 53 68 *************
104 30 69 *******
105 0 69
106 24 69 ******
107 27 70 ******
108 14 71 ****
109 24 71 ******
110 21 72 *****
111 25 72 ******
112 29 73 *******
113 26 74 ******
114 23 74 *****
115 11 75 **
116 12 75 ***
117 18 75 ****
118 13 76 ***
119 4 76 **
120 0 76
121 15 76 ***
122 20 77 *****
123 21 77 *****
124 16 78 ****
125 14 78 ***
126 20 78 *****
127 12 79 ***
128 0 79
129 16 79 ****
130 16 79 ****

Figure 50 (cont'd)

```
131    23 80 *****
132    20 80 *****
133    10 81 **
134     0 81
135    16 81 ****
136    13 81 ***
137    18 82 ****
138    12 82 ***
139    13 82 ***
140     0 82
141    17 83 ****
142    13 83 ***
143    14 84 ***
144    15 84 ***
145     0 84
146    14 84 ***
147    12 85 ***
148    19 85 ****
149     0 85
150    13 85 ***
151    11 86 **
152    17 86 ****
153     0 86
154    10 86 **
155    17 87 ****
156    14 87 ***
157     0 87
158    17 87 ****
159    24 88 ******
160     0 88
161    17 88 ****
162    17 89 ****
163    26 89 ******
164     0 89
165    21 90 *****
166    16 90 ****
167     0 90
168    13 91 ***
169     5 91 *
170     0 91
171     9 91 **
172     5 91 *
173     0 91
174    12 91 ***
175     4 92 *
176    10 92 **
177     3 92
178     5 92 *
179    12 92 ***
180     7 93 *
181    16 93 ****
182    10 93 **
183     6 93 *
184     9 94 **
185    12 94 ***
186    13 94 ***
187    10 94 **
```

Figure 50 (cont'd)

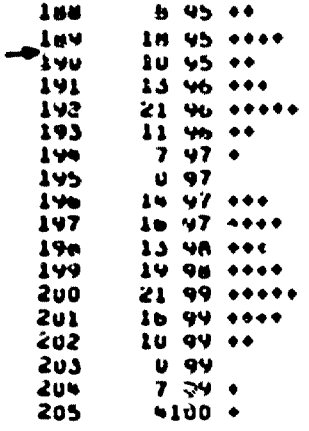


Figure 51

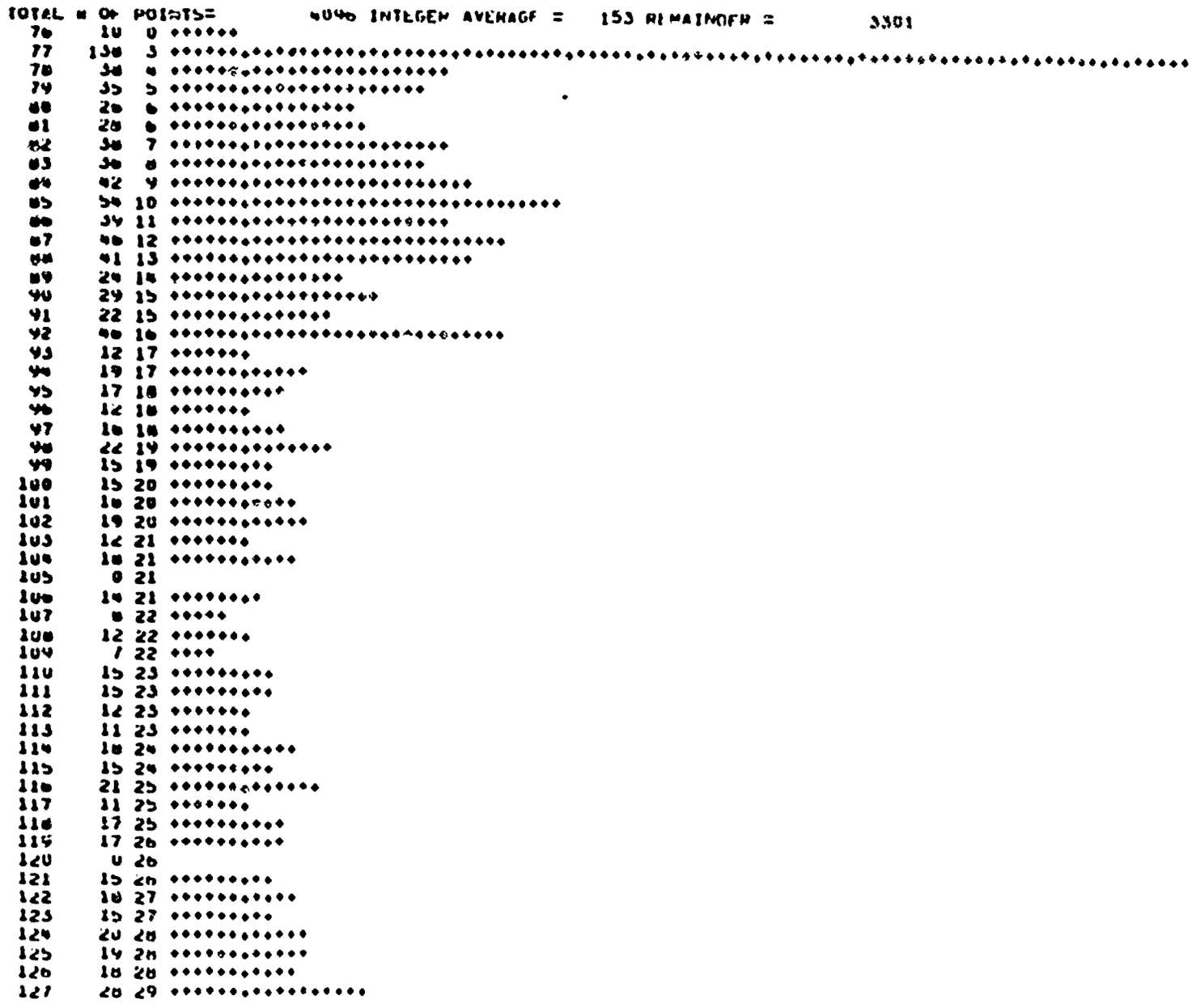

TOTAL # OF POINTS= %OF% INTEGER AVERAGE = 153 REMAINDER = 3301
76 10 0
77 138 3
78 38 4
79 35 5
80 26 6
81 28 6
82 38 7
83 36 8
84 42 9
85 54 10
86 34 11
87 46 12
88 41 13
89 24 14
90 29 15
91 22 15
92 46 16
93 12 17
94 19 17
95 17 18
96 12 18
97 16 18
98 22 19
99 15 19
100 15 20
101 16 20
102 19 20
103 12 21
104 18 21
105 0 21
106 14 21
107 8 22
108 12 22
109 7 22
110 15 23
111 15 23
112 12 23
113 11 23
114 18 24
115 15 24
116 21 25
117 11 25
118 17 25
119 17 26
120 0 26
121 15 26
122 18 27
123 15 27
124 20 28
125 19 28
126 18 28
127 28 29

Figure 51 (cont'd)

128	0	29
129	20	30
130	21	30
131	39	31
132	25	32
133	26	32
134	0	32
135	26	33
136	44	34
137	38	35
138	31	36
139	34	37
140	0	37
141	34	37
142	31	38
143	29	39
144	18	39
145	0	39
146	18	40
147	21	40
148	19	41
149	0	41
150	23	41
151	19	42
152	16	42
153	0	42
154	12	42
155	13	43
156	20	43
157	0	43
158	23	44
159	28	44
160	0	44
161	27	45
162	25	46
163	28	46
164	0	46
165	16	47
166	23	47
167	0	47
168	33	48
169	18	49
170	0	49
171	23	49
172	26	50
173	0	50
174	26	50
175	45	52
176	26	52
177	43	53
178	35	54
179	45	55
180	45	56
181	93	59
182	48	60
183	61	61
184	77	63

Figure 51 (cont'd)

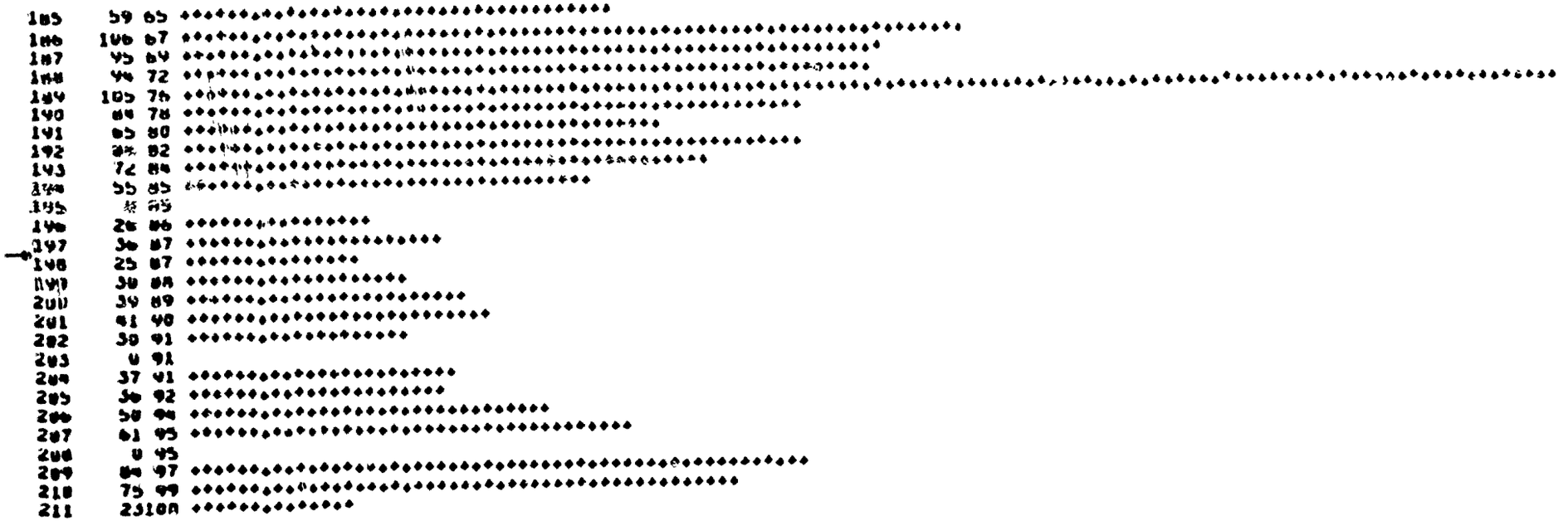


Figure 52

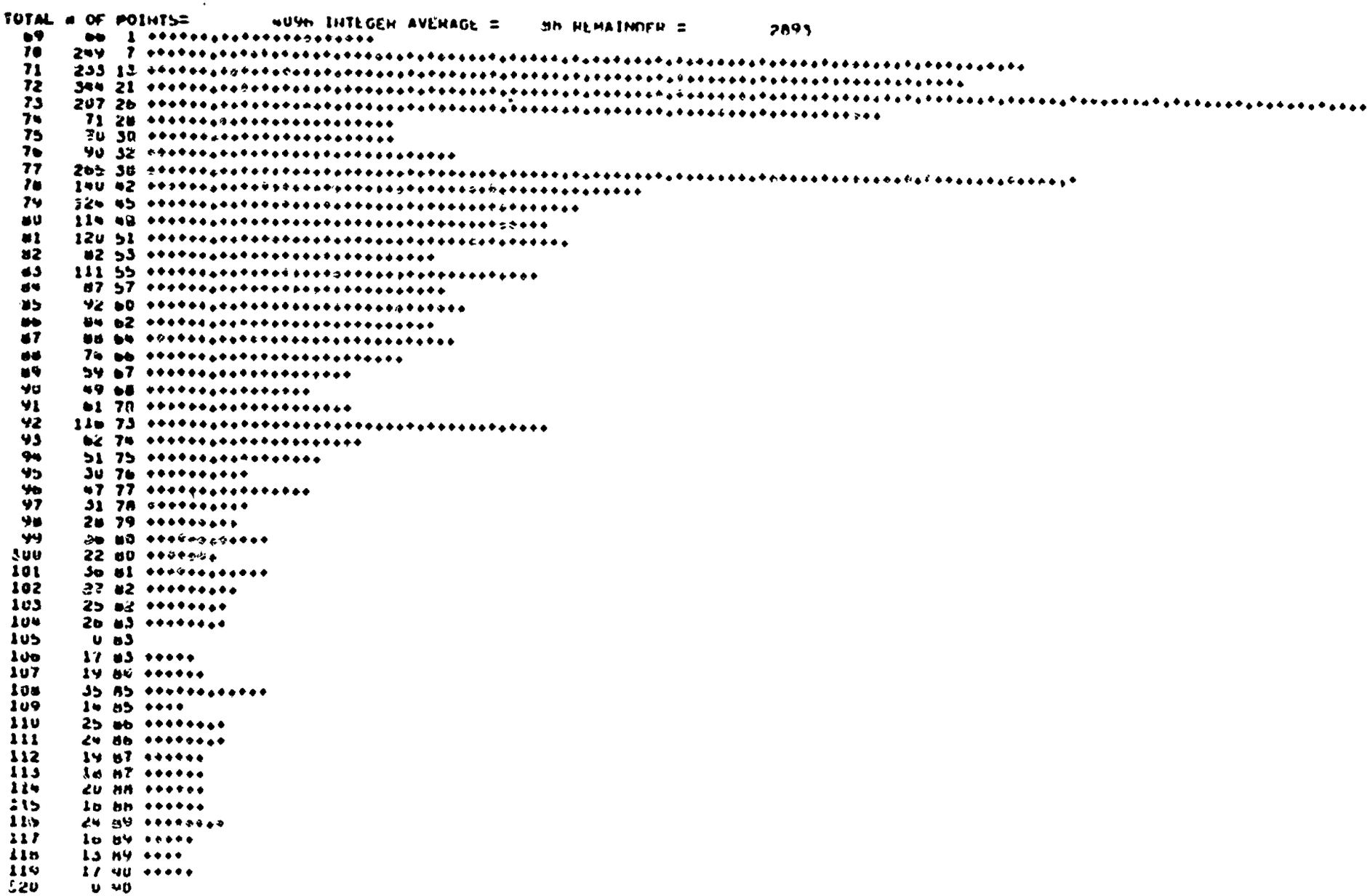


Figure 52 (cont'd)

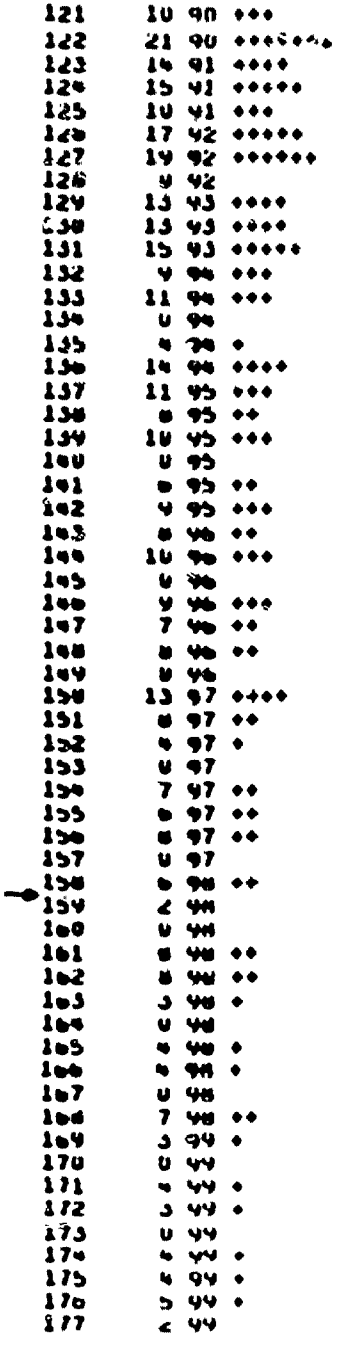


Figure 52 (cont'd)

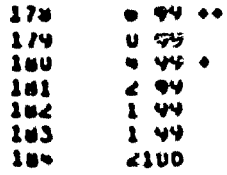


Figure 53

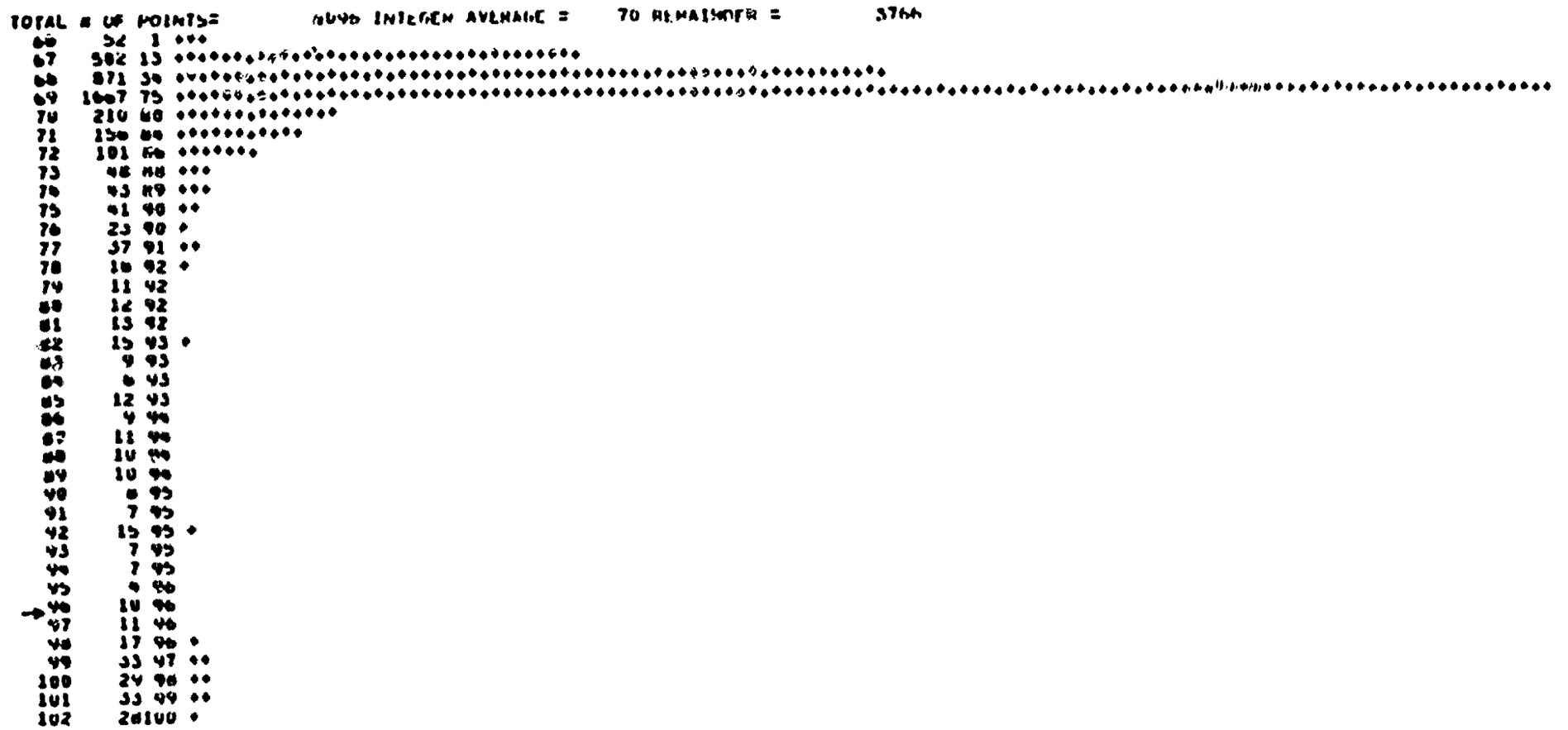


Figure 54

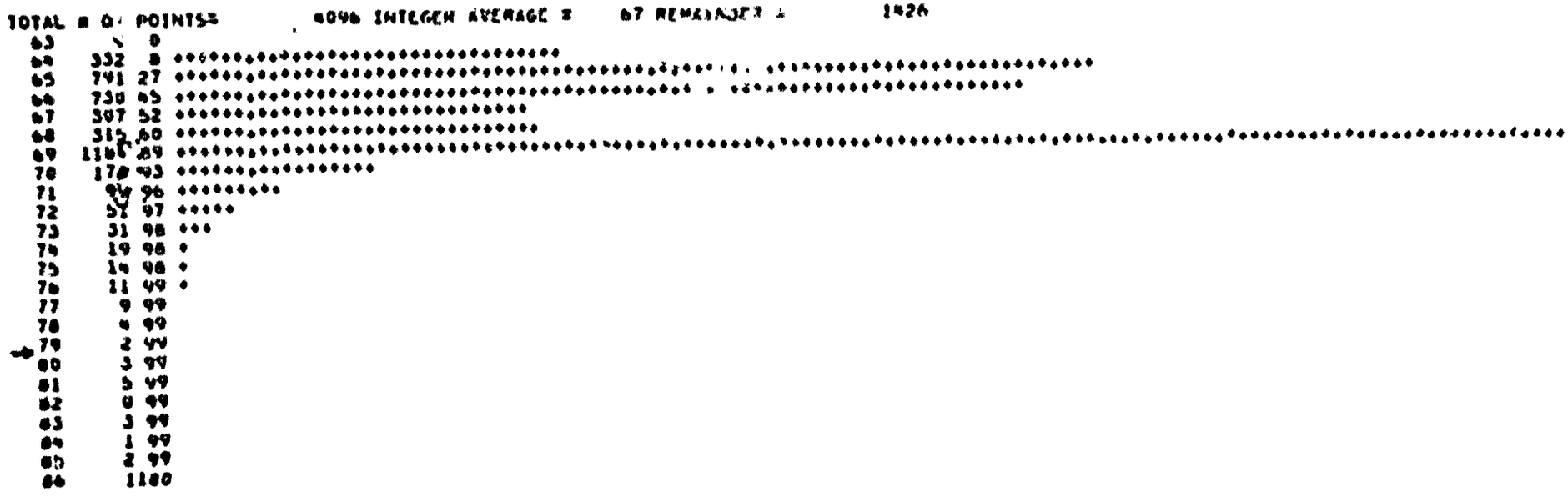

TOTAL # OF POINTS=
4096 INTEGER AVERAGE =
67 REMAINDER =
1426
63 0
64 332 8
65 791 27
66 750 45
67 307 52
68 315 60
69 1186 89
70 170 93
71 96 96
72 57 97
73 31 98
74 19 98
75 14 98
76 11 99
77 9 99
78 4 99
79 2 99
80 3 99
81 5 99
82 0 99
83 3 99
84 1 99
85 2 99
86 1180

Figure 55

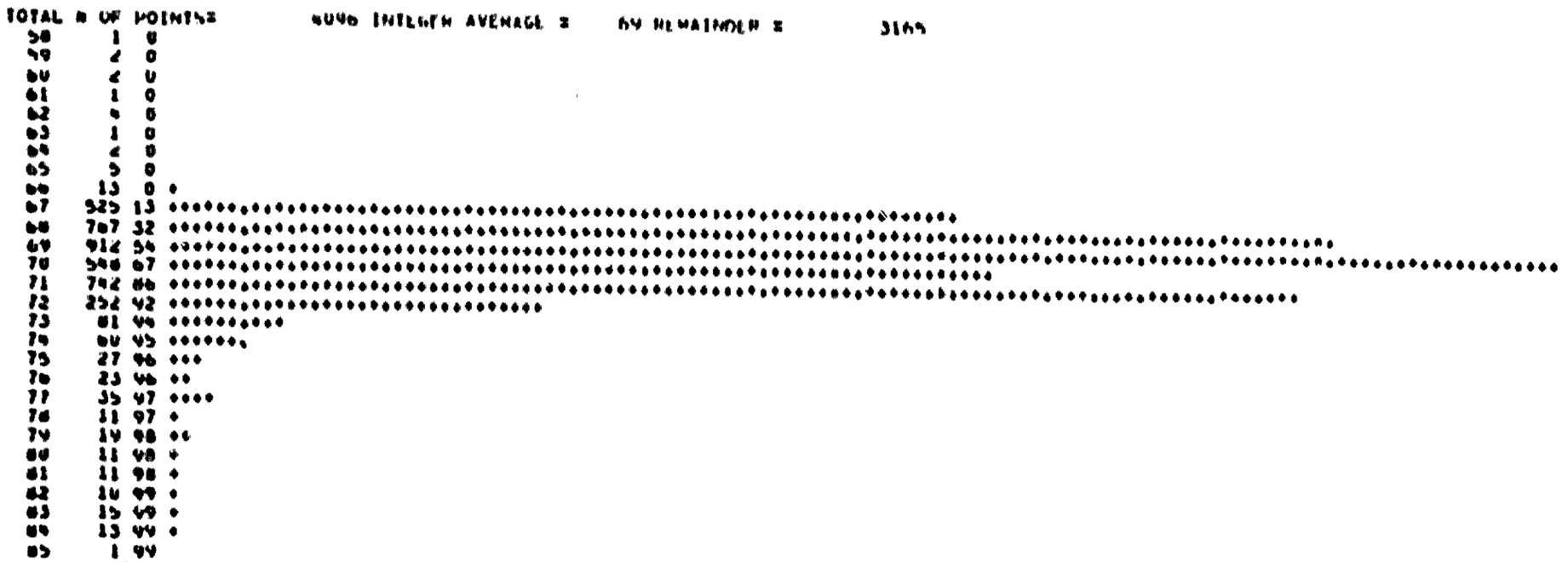

TOTAL # OF POINTS=
4096 INTERVAL AVERAGE =
BY REMAINDER =
3169
58 1 0
59 2 0
60 2 0
61 1 0
62 4 0
63 1 0
64 2 0
65 5 0
66 13 0 *
67 525 13 **
68 787 32 **
69 912 54 **
70 548 67 **
71 742 86 **
72 252 92 ********************************
73 81 94 **********
74 60 95 *******
75 27 96 ***
76 23 96 **
77 35 97 ****
78 11 97 *
79 14 98 **
80 11 98 *
81 11 98 *
82 14 99 *
83 15 99 *
84 13 99 *
85 1 99
86 2100

Figure 55

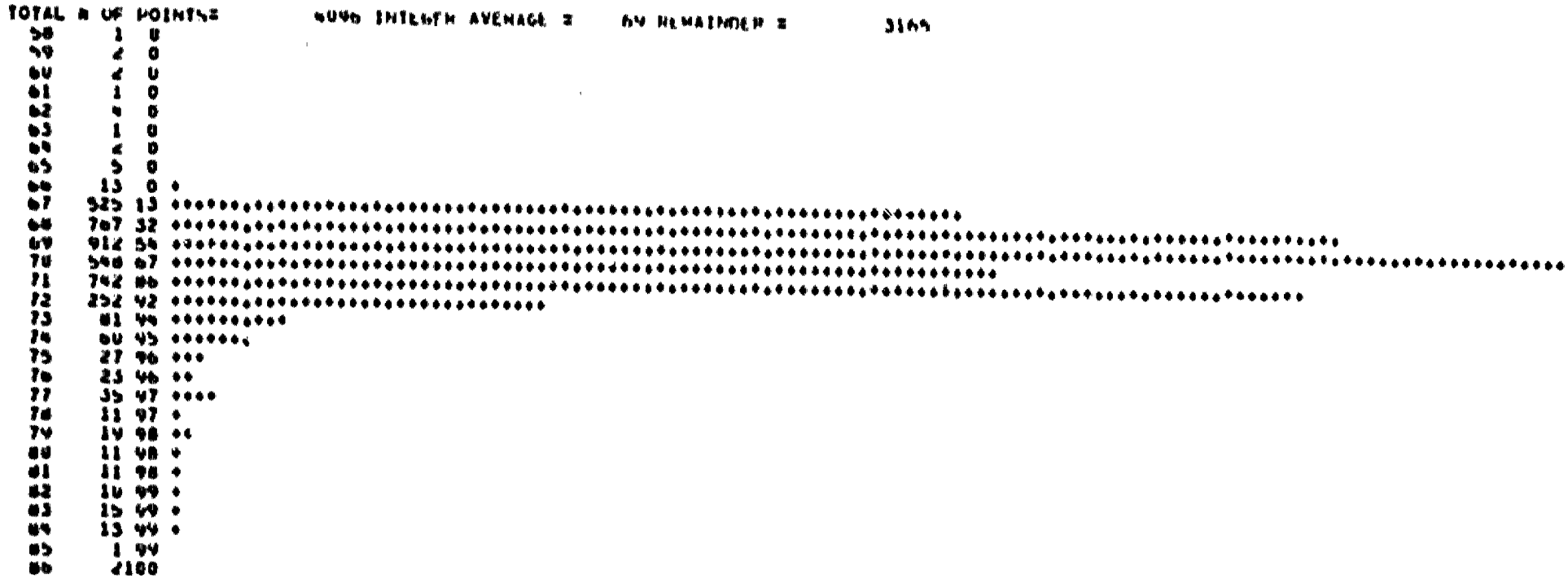

TOTAL # OF POINTS=
6096 INTEGER AVERAGE =
BY REMAINDER =
3169
58 1 0
59 2 0
60 2 0
61 1 0
62 4 0
63 1 0
64 2 0
65 5 0
66 13 0 *
67 525 13 **
68 787 32 **
69 912 54 **
70 548 87 **
71 742 86 **
72 252 92 ********************************
73 81 96 **********
74 60 95 *******
75 27 96 ***
76 23 96 **
77 35 97 ****
78 11 97 *
79 19 98 **
80 11 98 *
81 11 98 *
82 10 99 *
83 15 99 *
84 13 99 *
85 1 99
86 2100

Figure 56

TOTAL # OF POINTS=   4094 INTEGER AVERAGE =   81 REMAINDER =   657

68	167	6
69	1050	29
70	397	39
71	373	48
72	275	55
73	135	58
74	98	61
75	89	63
76	75	65
77	135	68
78	51	69
79	44	70
80	49	71
81	38	72
82	31	73
83	35	74
84	49	75
85	36	76
86	35	77
87	39	78
88	47	79
89	38	80
90	39	81
91	31	82
92	75	83
93	35	84
94	20	85
95	20	85
96	25	86
97	16	86
98	22	87
99	17	87
100	14	87
101	17	88
102	17	88
103	12	89
104	15	89
105	0	89
106	13	89
107	13	90
108	13	90
109	0	90
110	12	90
111	11	91
112	10	91
113	11	91
114	0	91
115	11	92
116	0	92
117	4	92
118	11	92
119	[illegible]	[illegible]

Figure 56 (cont'd)

120	0	92	
121	9	93	*
122	9	93	*
123	9	93	*
124	9	93	*
125	5	93	
126	0	93	
127	7	94	
128	0	94	
129	6	94	
130	10	94	*
131	5	94	
132	9	94	*
133	9	95	*
134	0	95	
135	8	95	
136	10	95	*
137	8	95	
138	9	95	*
139	10	96	*
140	0	96	
141	10	96	*
142	8	96	
143	6	96	
144	8	96	
145	8	96	
146	5	97	
147	5	97	
148	3	97	
149	0	97	
150	32	97	*
151	10	97	*
152	8	97	
153	0	97	
154	2	97	
155	8	98	
156	7	98	
157	11	98	
158	5	98	
159	9	98	*
160	0	98	
161	8	98	
162	5	98	
163	2	98	
164	0	98	
→ 165	3	98	
166	3	99	
167	0	99	
168	7	99	
169	2	99	
170	0	99	
171	6	99	
172	8	99	
173	0	99	
174	8	99	
175	1	99	
176	5	99	

Figure 56 (cont'd)

```
177     2 99
178     0 99
179     3 99
180     3 99
181     1100
```

Figure 57

Figure 57 (cont'd)

```
120      0 91
121      7 92 *
122     11 92 **
123      5 92 *
124      8 92 *
125      3 92
126      6 93 *
127      4 93
128      0 93
129     13 93 **
130      6 93 *
131      7 93 *
132     11 94 **
133      6 94 *
134      0 94
135      4 94
136     15 94 ***
137      8 94 *
138      7 95 *
139     11 95 **
140      0 95
141      2 95
142      8 95 *
143      8 95 *
144      3 95
145      0 95
146      6 96 *
147      6 96 *
148      7 96 *
149      0 96
150     10 96 **
151      8 96 *
152      4 96
153      0 96
154      7 97 *
155      8 97 *
156      6 97 *
157      0 97
158      8 97 *
159      3 97
160      0 97
161     12 98 **
162     12 98 **
163      5 98 *
164      0 98
165      7 98 *
166      4 98
167      0 98
168      5 98 *
169      5 98 *
170      0 98
171      5 99 *
172      2 99
173      0 99
174      8 99 *
175      3 99
176      1 99
```

Figure 57 (cont'd)

```
177     6 99 *
178     6 99
179     2 99
180     2 99
181     7 99 *
182     3 99
183     2 99
184     1100
```

Figure 58

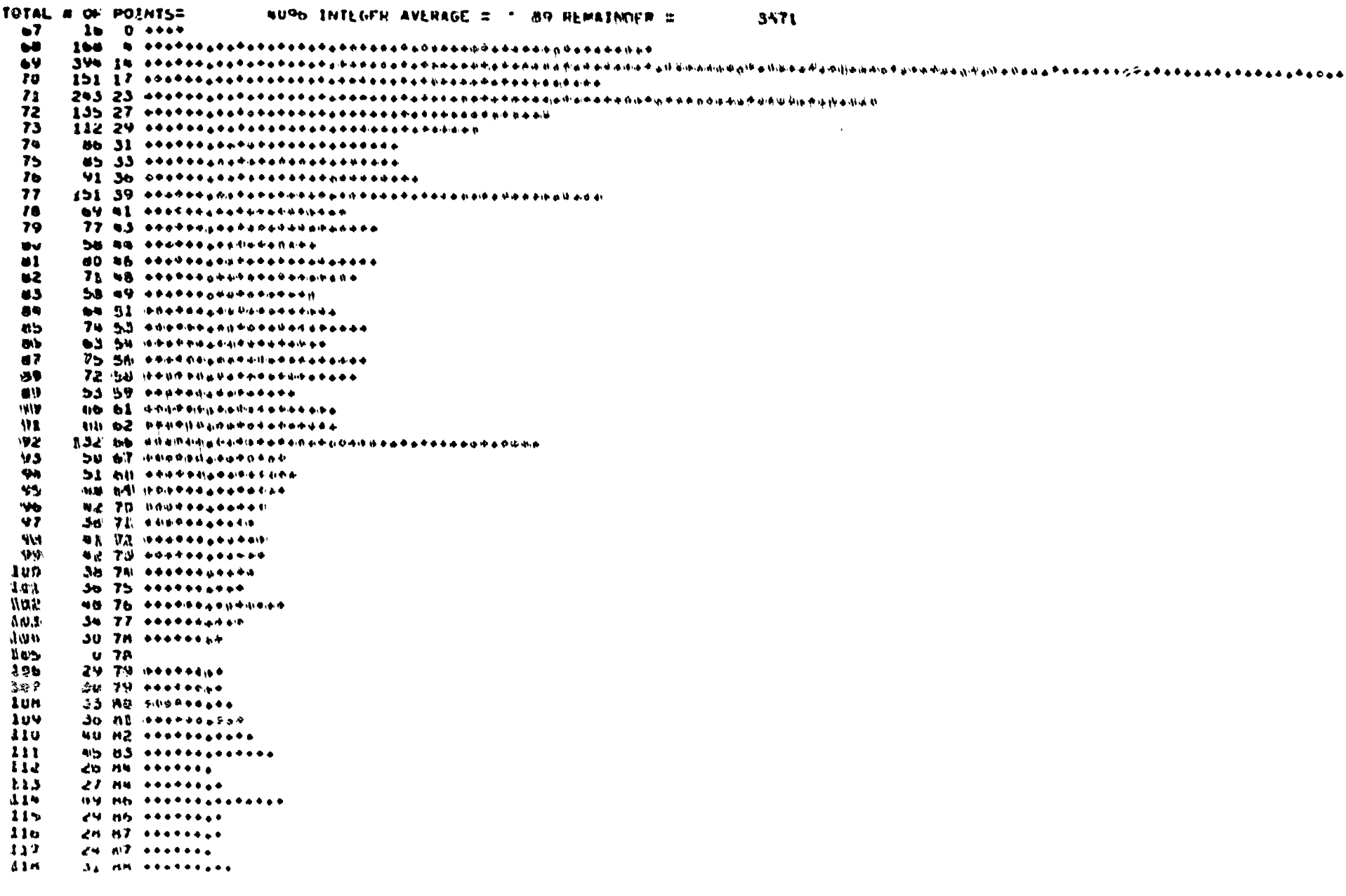


Figure 58 (cont'd)

```
119    24 89 ******+
120     0 89
121    25 89 ******+
122    20 90 +++++
123    13 90 +++
124    18 91 +++++
125    17 91 +++++
126    21 92 ******+
127    15 92 ++++
128     0 92
129    18 92 +++++
130    19 93 +++++
131    23 93 ******+
132     4 94 ++
133    11 94 +++
134     0 94
135    18 94 +++++
136    19 95 +++++
137    11 95 +++
138    20 96 +++++
139    16 96 ++++
140     0 96
141     8 96 ++
142    10 96 ++
143     5 96
144     9 97 ++
145     0 97
146     7 97 ++
147     3 97 +
148    11 97 +++
149     0 97
150     5 97 +
151    11 98 +++
152     4 98 +
153     0 98
154     8 98 ++
155     7 98 ++
156     6 98 ++
157     0 98
158    13 99 +++
159     3 99 +
160     0 99
161     4 99 +
162     5 99 +
163     2 99
164     0 99
165     4 99 +
166     4 99 +
167     0 99
168     1 99
169     2 99
170     0 99
171     3 99
172     0 99
173     0 99
174     2 99
175     0 99
```

Figure 58 (cont'd)

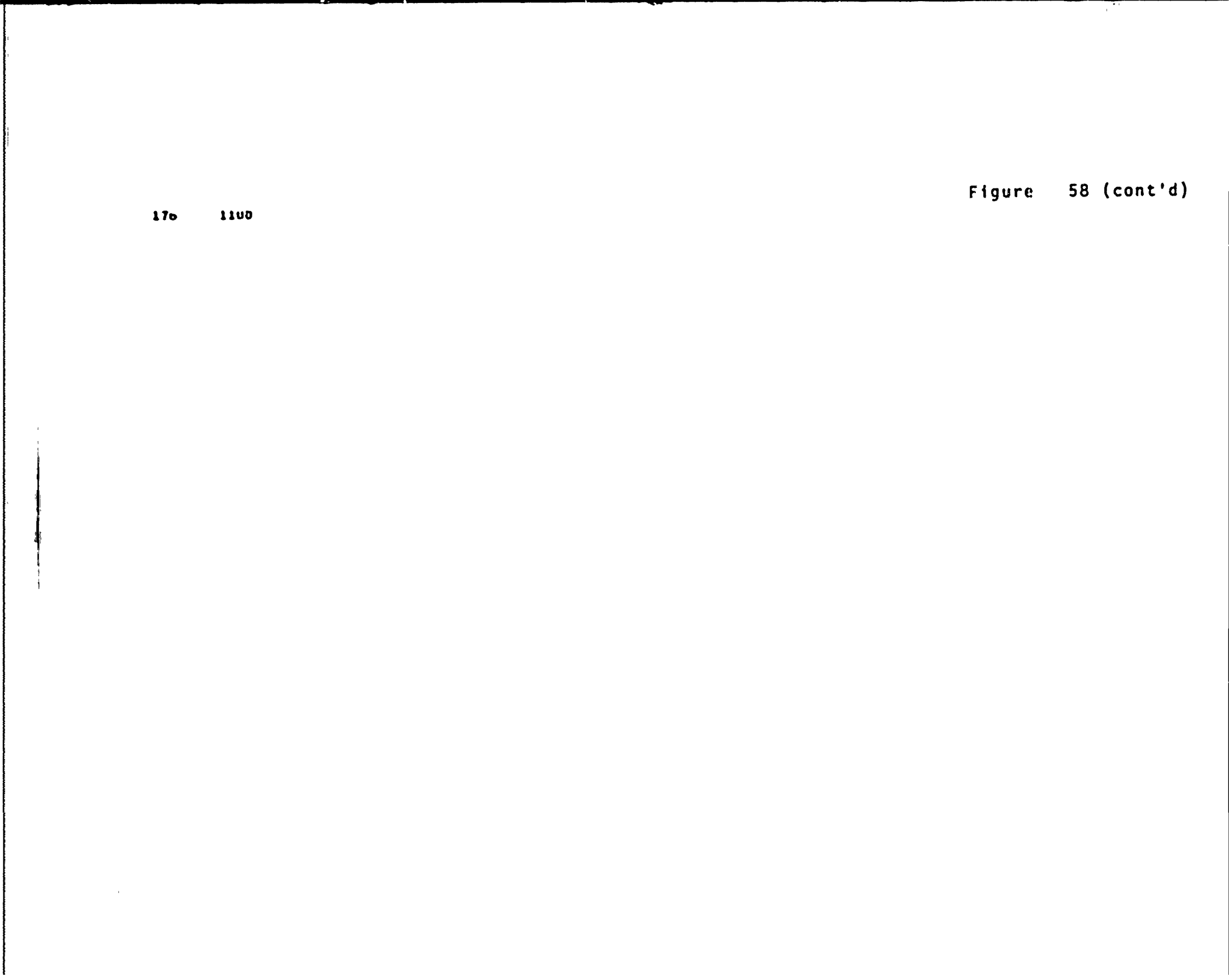


Figure 59

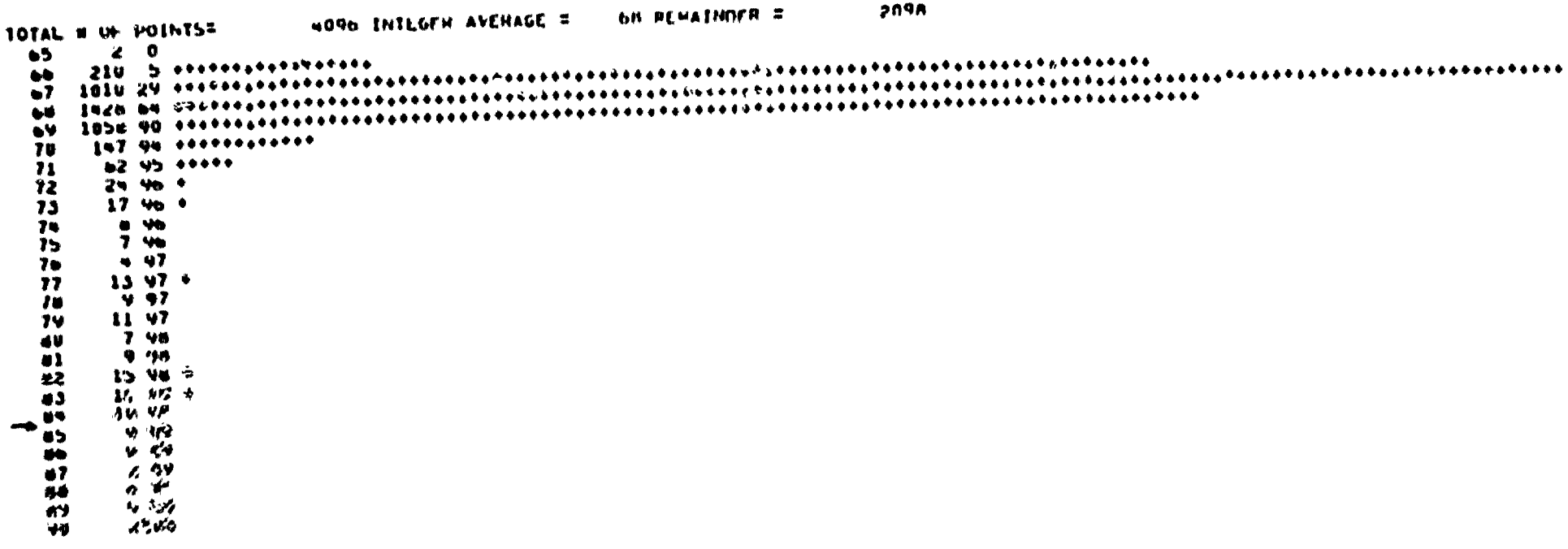
TOTAL # OF POINTS= 4096 INTEGER AVERAGE = 76 REMAINDER = 3124

67	54	1
68	265	7
69	1215	37
70	413	47
71	294	54
72	190	59
73	131	62
74	138	65
75	95	68
76	66	69
77	138	73
78	66	74
79	40	75
80	60	77
81	46	78
82	44	79
83	35	80
84	35	81
85	32	82
86	44	83
87	41	84
88	60	85
89	43	86
90	34	87
91	36	88
92	85	90
93	33	91
94	40	92
95	21	92
96	20	93
97	12	93
98	12	93
99	9	94
100	15	94
101	13	94
102	9	94
103	7	95
104	8	95
105	6	95
106	7	95
107	8	95
108	8	95
109	9	96
110	6	96
111	9	96
112	6	96
113	14	97
114	5	97
115	7	97
116	4	97
117	4	97
118	3	97

Figure 59 (cont'd)

```
119     3 97
120     0 97
121     1 97
122     4 97
123     5 97
124     2 97
125     2 97
126     1 98
127     3 98
128     0 98
129     0 98
130     5 98
131     5 98
132     3 98
133     0 98
134     0 98
135     0 98
136     0 98
137     0 99
138     0 99
139     0 99
140     0 99
141     0 99
142     0 99
143     0 99
144     3 99
145     0 99
146     3 99
147     2 99
148     3 99
149     0 99
150     0 99
151     0 99
152     1 99
153     0 99
154     1 99
155     1 99
156     0 99
157     0 99
158     1100
```

Figure 60



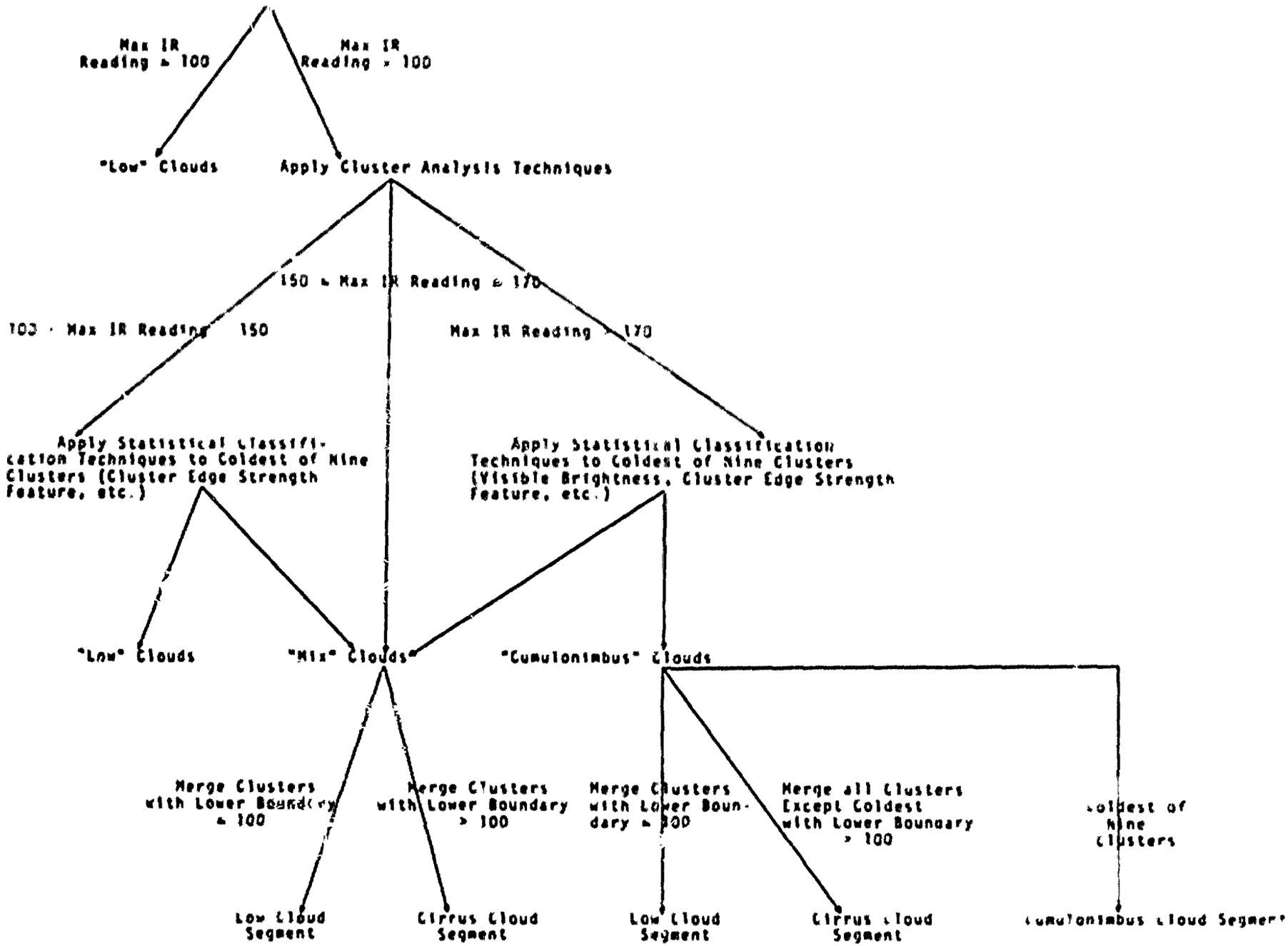


Figure 61 Decision Procedure for Cluster Edge Strength Model

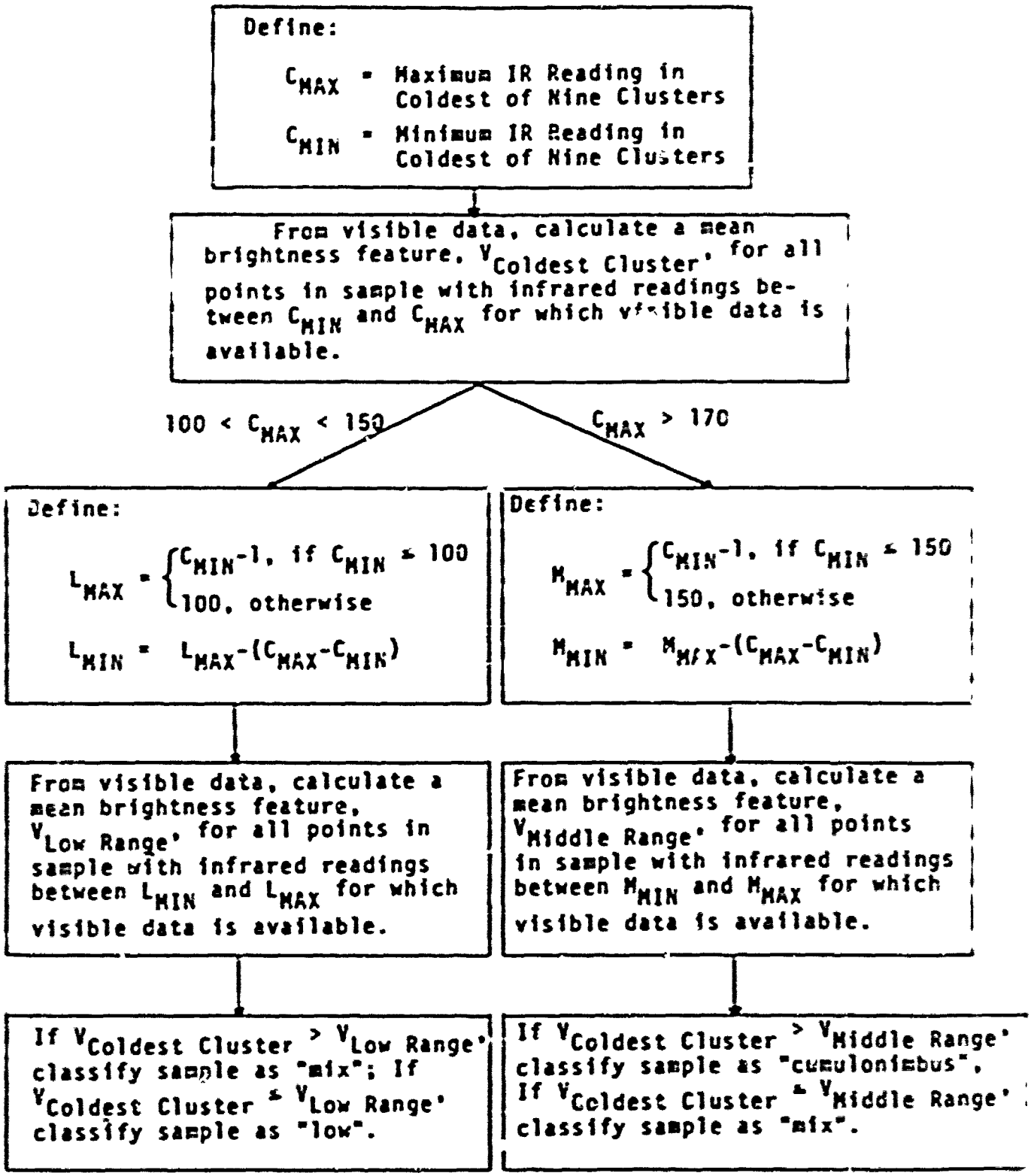


Figure 52. Algorithm for Classification of Cloud Samples at Stage 3 of Decision Procedure given in Figure 61 Using Visible Brightness Feature.

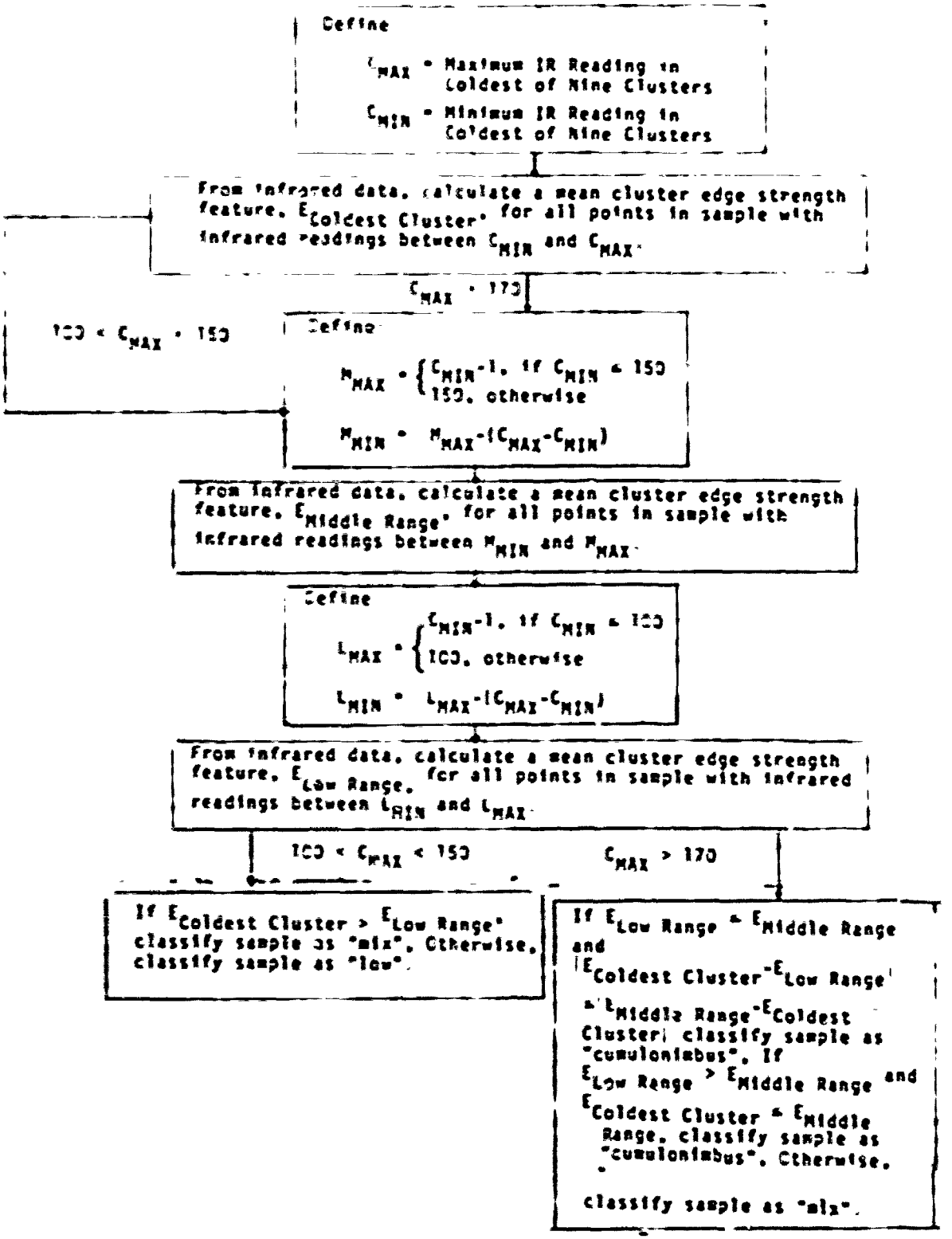


Figure 63 Algorithm for Classification of Cloud Samples at Stage 3 of Decision Procedure given in Figure 61 Using Cluster Edge Strength Feature

Figure 64 Original and Segmented Images for Sample Numbers 1-16.

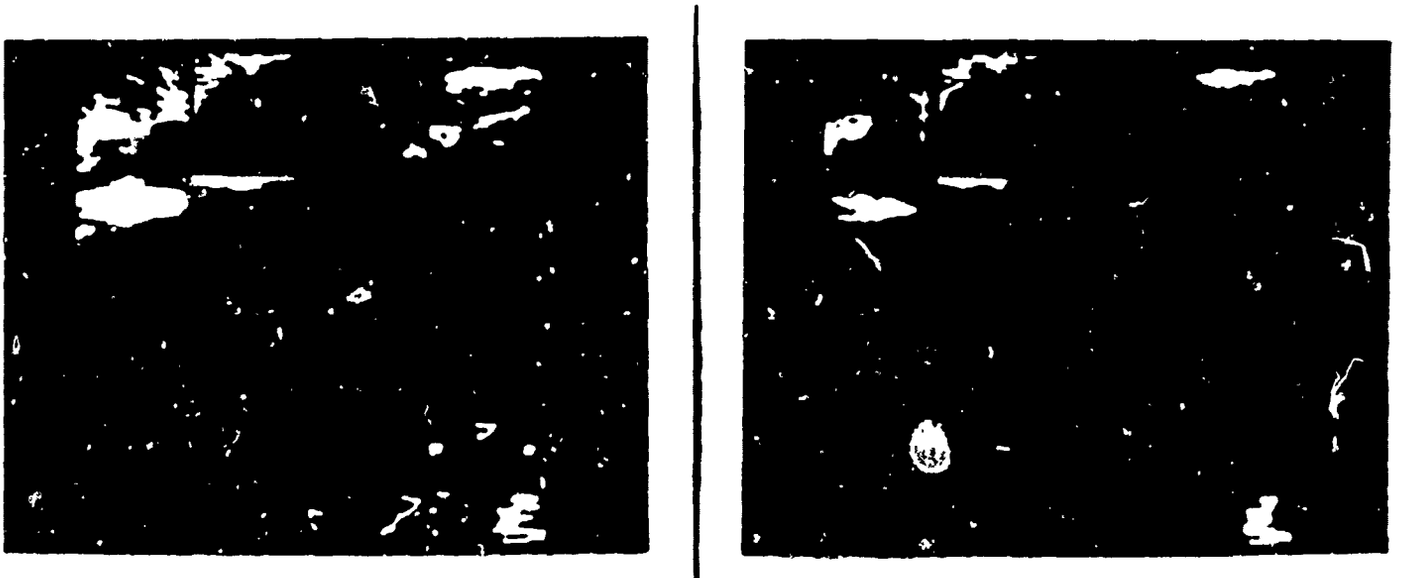


Figure 65. Original and Segmented Images for Sample Numbers 17-32.

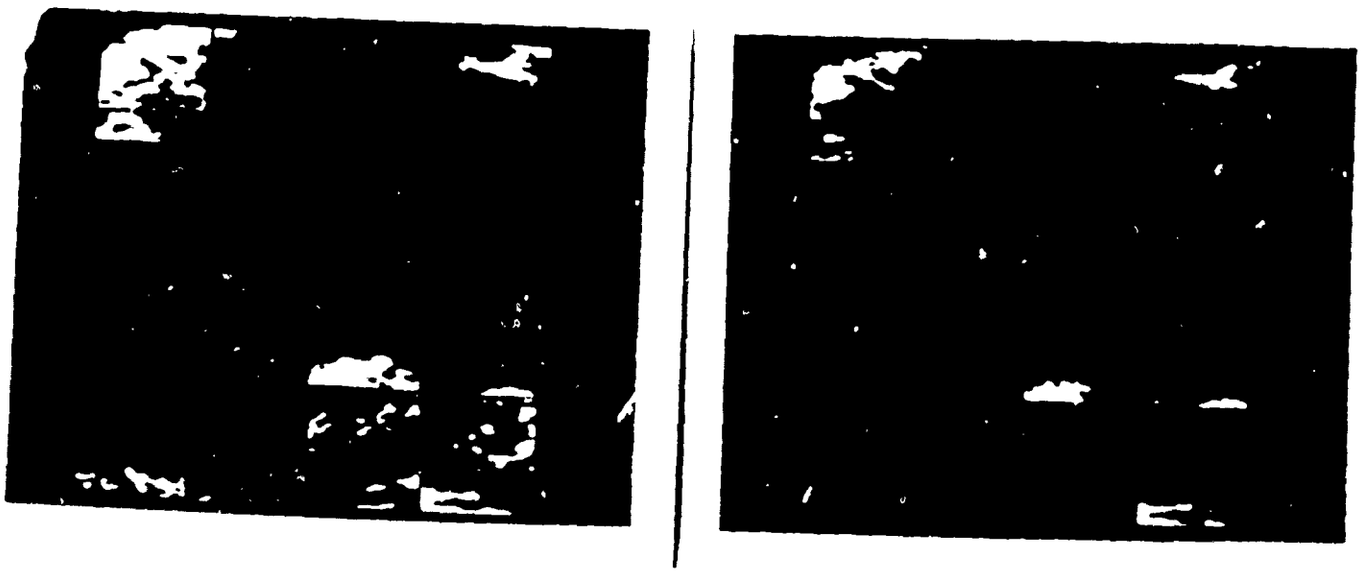


Figure 6. Orignal and Segmented Images for Sample Numbers 33-48.

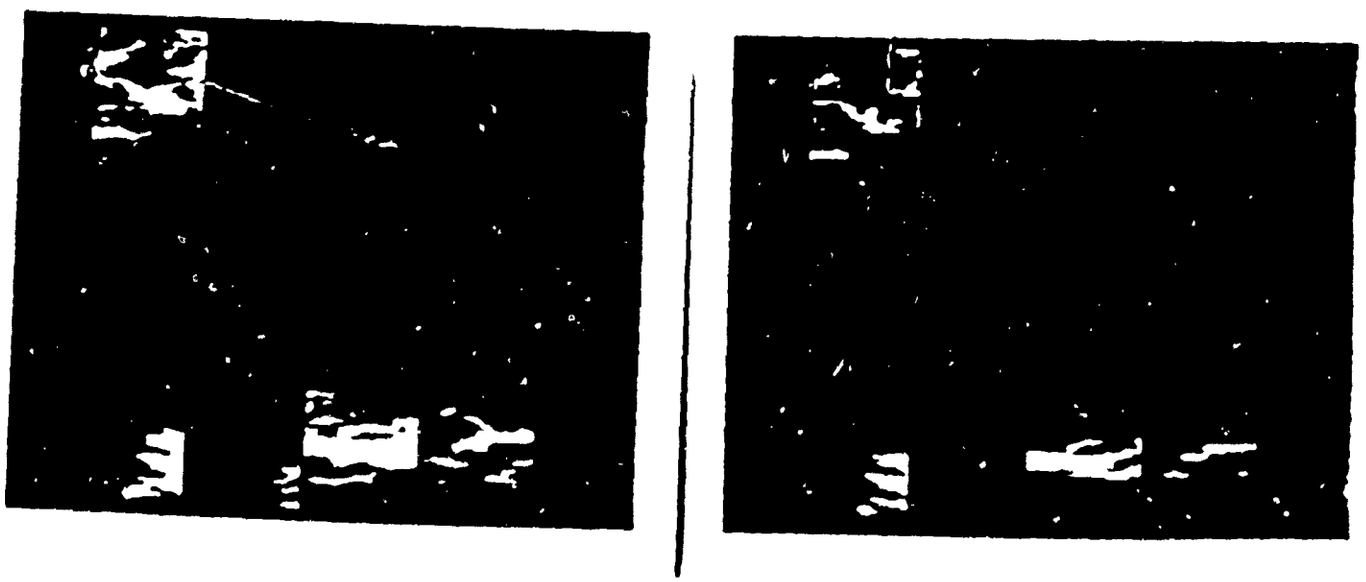


Figure 6 Original and segmented images for Sample Numbers 49-64.

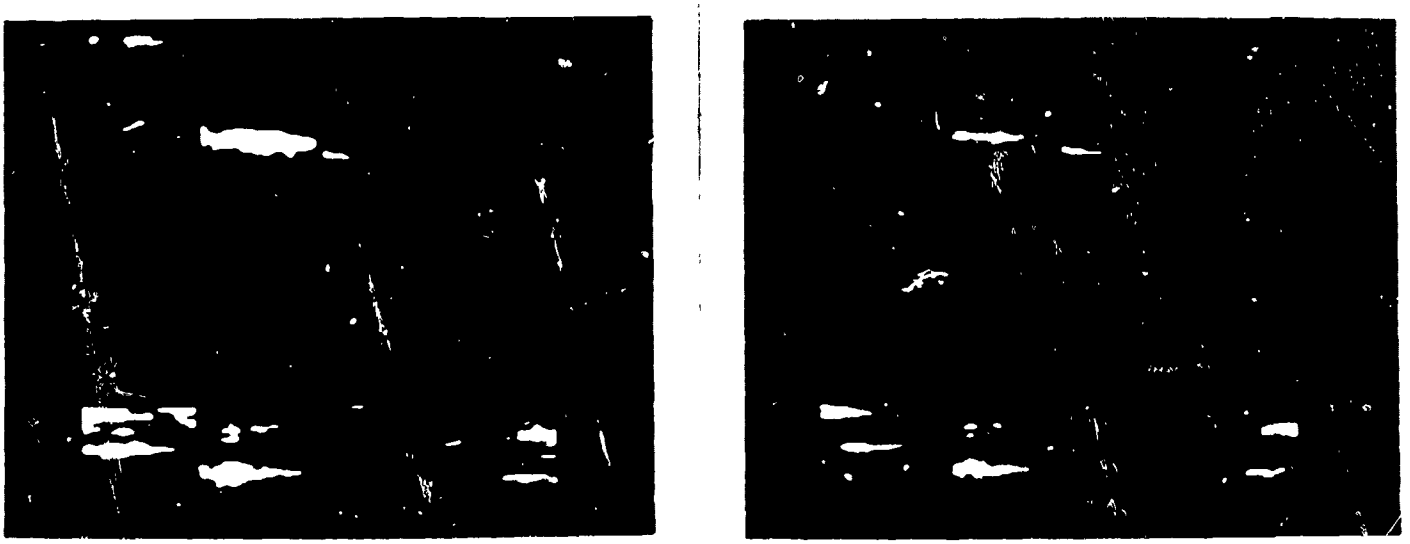


Figure 6 Original and Segmented Images for Sample Numbers 65-80.

Figure 69 Original and Segmented Images of Sample Numbers 81-96.

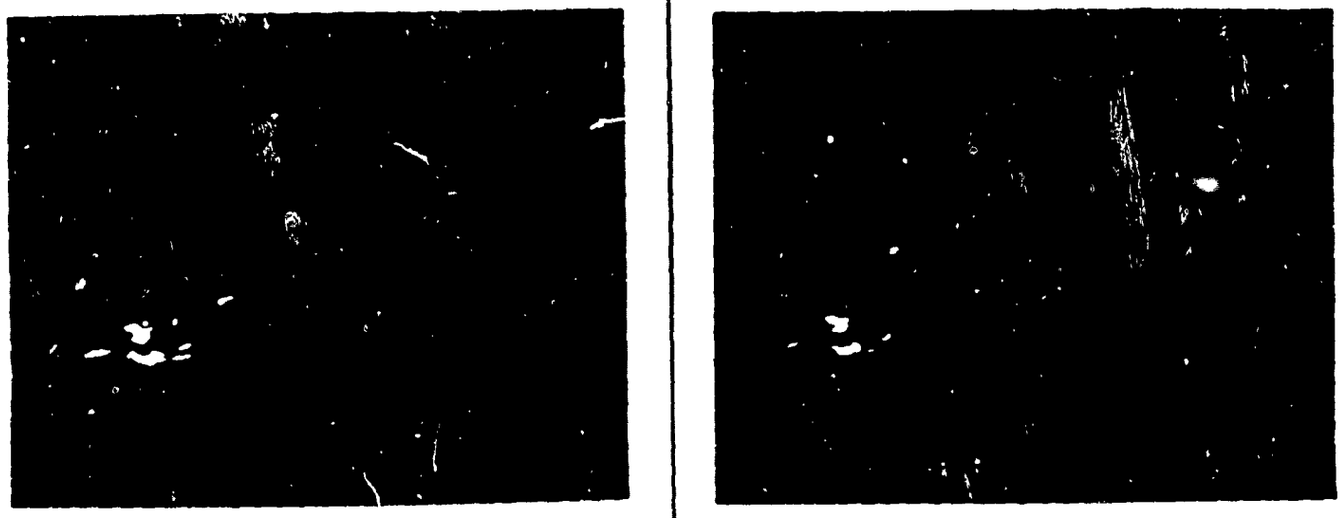


Figure 7[illegible] Original and Segmented Images for Sample Numbers 97-107.

SAMPLE NUMBER	CLOUD TYPE CLASSIFICATION	GEOGRAPHICAL LOCATION* LATITUDE, LONGITUDE
1	Cumulonimbus	12.5N, 122.5W
2	Cumulonimbus	12.5N, 117.5W
3	Cumulonimbus	12.5N, 112.5W
4	Mix	12.5N, 107.5W
5	Low (Middle)	12.5N, 102.5W
6	Low	12.5N, 97.5W
7	Low	12.5N, 92.5W
8	Mix	12.5N, 82.5W
9	Mix	12.5N, 77.5W
10	Mix	12.5N, 67.5W
11	Mix	12.5N, 62.5W
12	Low	12.5N, 57.5W

*Geographical location of the center point of the data samples can be used to locate the data samples in the visible and infrared satellite images of Figures 1 and 2 of [1].

Table 1. Classification and Location of Data Samples for Pilot Segmentation Study.

SAMPLE NUMBER	DESCRIPTION OF COLDEST OF NINE CLUSTERS OF INFRARED HISTOGRAM			
	OBSERVATIONS	MEAN	STANDARD DEVIATION	PARTITION SIZES*
1	190-206	199.35	4.32	9,10
2	190-205	196.87	4.14	9,10
3	198-211	205.34	4.01	7,8,9,10
4	159-184	170.27	7.14	8,9
5	97-102	99.93	1.51	8,9
6	80-86	82.27	1.95	9,10
7	80-86	82.30	1.58	7,8,9
8	165-181	172.36	4.67	8,9,10
9	168-184	175.38	4.88	9,10
10	150-176	158.19	6.23	9,10
11	135-158	141.30	5.50	9,10
12	85-90	86.86	1.68	7,8,9

*The partition sizes correspond to those values of K ≤ 10 (the number of clusters) for which the coldest cluster was identical to the coldest cluster for K=9.

Table 2. Analysis of Coldest Cluster of Data Samples for Pilot Segmentation Study Where Number of Clusters K=9.

SAMPLE NUMBER	VISIBLE BRIGHTNESS FEATURE ANALYSIS			CLUSTER EDGE STRENGTH FEATURE ANALYSIS			DECISION PROCEDURE CLASSIFICATION		CLOUD TYPE
	Coldest Cluster	Middle Range	Low Range	Coldest Cluster	Middle Range	Low Range	Brightness Feature	Edge Feature	
1	32.27	27.15		5.67	21.16	3.15	Cumulonimbus	Cumulonimbus	Cumulonimbus
2	31.80	22.32		5.50	15.73	4.05	Cumulonimbus	Cumulonimbus	Cumulonimbus
3	42.53	37.93		3.06	11.39	8.30	Cumulonimbus	Cumulonimbus	Cumulonimbus
4	30.59	27.07		12.17	12.93	4.83	Cumulonimbus	Mix	Mix
5	34.63		27.56	3.18		8.12	Mix	Low	Low
6							Low	Low	Low
7							Low	Low	Low
8	22.30	23.36		25.05	25.21	11.11	Mix	Mix	Mix
9	NO VISIBLE DATA AVAILABLE CORRESPONDING TO COLDEST CLUSTER POINTS			23.00	26.04	14.74	---	Mix	Mix
10	25.05	22.94		17.92	19.22	9.80	Cumulonimbus	Mix	Mix
11							Mix	Mix	Mix
12							Low	Low	Low

Table 3. Classification Results for Samples in Pilot Study Using Decision Procedure Approach of Figure 61.

Table 4. Classification results using cluster edge strength model.

SAMPLE NUMBER	MAXIMUM IR OBSERVATION	COLDEST CLUSTER		MIDDLE RANGE		LOW RANGE		DECISION PROCEDURE CLASSIFICATION	CLOUD-TRUTH CLASSIFICATION
		OBSERVATIONS	EDGE STRENGTH	OBSERVATIONS	EDGE STRENGTH	OBSERVATIONS	EDGE STRENGTH		
1	206	190-206	5.67	134-150	21.16	84-100	3.15	Cumulonimbus	Cumulonimbus
2	205	190-205	5.50	135-150	15.73	85-100	4.05	Cumulonimbus	Cumulonimbus
3	211	198-211	3.06	137-150	11.39	87-100	8.30	Cumulonimbus	Cumulonimbus
4	184	159-184	12.17	125-150	12.93	75-100	4.83	Mix	Mix
5	102	97-102	3.18			91-96	8.12	Low	Low (Middle)
6	86							Low	Low
7	86							Low	Low
8	181	165-181	25.05	134-150	25.21	84-100	11.11	Mix	Mix
9	184	168-184	23.00	134-150	26.04	84-100	14.74	Mix	Mix
10	176	150-176	17.92	123.149	19.22	74-100	9.80	Mix	Mix
11	158							Mix	Mix
12	90							Low	Low
13	79							Low	Low
14	89							Low	Low
15	162							Mix	Mix
16	176	161-176	6.24	135-150	13.70	85-100	11.99	Cumulonimbus	Cumulonimbus
17	192	177-192	4.39	135-150	13.52	85-100	8.41	Cumulonimbus	Cumulonimbus
18	181	165-181	7.40	134-150	16.03	84-100	7.52	Cumulonimbus	Cumulonimbus
19	186	162-186	11.30	126-150	13.05	76-100	3.55	Mix	Cumulonimbus
20	206	190-206	4.77	134-150	14.32	84-100	5.94	Cumulonimbus	Cumulonimbus

Table 4 (cont'd)

SAMPLE NUMBER	MAXIMUM IR OBSERVATION	COLDEST CLUSTER		MIDDLE RANGE		LOW RANGE		DECISION PROCEDURE	CLOUD-TRUTH
		OBSERVATIONS	EDGE STRENGTH	OBSERVATIONS	EDGE STRENGTH	OBSERVATIONS	EDGE STRENGTH	CLASSIFICATION	CLASSIFICATION
21	210	198-210	3.43	138-150	14.34	88-100	6.29	Cumulonimbus	Cumulonimbus
22	201	187-201	5.96	136-150	19.85	86-100	8.52	Cumulonimbus	Cumulonimbus
23	99							Low	Low
24	78							Low	Low
25	81							Low	Low
26	77							Low	Low
27	136	115-136	24.25			79-100	7.70	Mix	Mix
28	143	123-143	10.87			80-100	5.47	Mix	Mix
29	116	101-116	22.25			85-100	8.86	Mix	Mix
30	154							Mix	Mix
31	180	161-180	12.01	131-150	12.37	81-100	5.35	Mix	Mix
32	196	178-196	4.72	132-150	10.65	82-100	13.21	Cumulonimbus	Cumulonimbus
33	192	180-192	4.18	138-150	14.18	88-100	10.29	Cumulonimbus	Cumulonimbus
34	172	150-172	8.49	127-149	16.25	78-100	4.03	Cumulonimbus	Cumulonimbus
35	121	106-121	9.52			85-100	6.57	Mix	Mix
36	191	174-191	6 54	133-150	15.83	83-100	3 71	Cumulonimbus	Cumulonimbus
37	150							Mix	Mix
38	133	116-133	13 25			83-100	3 29	Mix	Mix
39	119	112-119	4.84			93-100	4 48	Mix	Mix
40	85							Low	Low
41	109	101-109	6 69			92-100	7.76	Low	Low (Middle)

Table 4 (cont'd)

SAMPLE NUMBER	MAXIMUM IR OBSERVATION	COLDEST CLUSTER		MIDDLE RANGE		LOW RANGE		DECISION PROCEDURE CLASSIFICATION	CLOUD-TRUTH CLASSIFICATION
		OBSERVATIONS	EDGE STRENGTH	OBSERVATIONS	EDGE STRENGTH	OBSERVATIONS	EDGE STRENGTH		
42	113	106-113	3.53			93-100	4.74	Low	Low (Middle)
43	194	183-194	2.48	139-150	15.16	89-100	7.68	Cumulonimbus	Cumulonimbus
44	193	174-193	5.46	131-150	26.94	81-100	8.12	Cumulonimbus	Cumulonimbus
45	181	158-181	9.26	127-150	7.98	77-100	4.06	Mix	Mix
46	174	152-174	7.71	128-150	5.49	78-100	11.70	Mix	Mix
47	192	166-192	12.58	124-150	10.52	74-100	5.94	Mix	Mix
48	191	178-191	5.47	137-150	8.16	87-100	9.94	Cumulonimbus	Cumulonimbus
49	197	174-197	9.44	127-150	13.65	77-100	7.69	Cumulonimbus	Cumulonimbus
50	132	121-132	22.65			89-100	12.37	Mix	Mix
51	92							Low	Low
52	96							Low	Low
53	90							Low	Low
54	90							Low	Low
55	104	97-104	3.51			89-96	3.82	Low	Low
56	88							Low	Low
57	93							Low	Low
58	116	111-116	2.48			95-100	5.77	Low	Low
59	174	144-174	17.00	113-143	2.89	73-100	3.94	Mix	Cumulonimbus
60	136	121-136	13.86			85-100	3.51	Mix	Mix
61	200	186-200	5.51	136-150	10.01	86-100	6.92	Cumulonimbus	Cumulonimbus
62	199	172-199	20.51	123-150	15.41	73-100	3.73	Mix	Mix
63	200	187-200	3.49	137-150	9 52	87-100	8 01	Cumulonimbus	Cumulonimbus
64	201	183-201	12.82	132-150	21.35	82-100	9 52	Cumulonimbus	Cumulonimbus
65	183	165-183	20.36	132-150	22 67	82-100	8 93	Mix	Mix
66	206	196-206	4.75	140-150	29.23	90-100	31 56	Cumulonimbus	Cumulonimbus
67	181	161-181	16 31	130-150	27 28	80-100	10 97	Cumulonimbus	Cumulonimbus

Table 4 (cont'd)

SAMPLE NUMBER	MAXIMUM IR OBSERVATION	COLDEST CLUSTER		MIDDLE RANGE		LOW RANGE		DECISION PROCEDURE	CLOUD-TRUTH
		OBSERVATIONS	EDGE STRENGTH	OBSERVATIONS	EDGE STRENGTH	OBSERVATIONS	EDGE STRENGTH	CLASSIFICATION	CLASSIFICATION
68	109	101-109	8.81			92-100	2.26	Mix	Mix
69	113	106-113	5.22			93-100	2.39	Mix	Mix
70	92							Low	Low
71	89							Low	Low
72	93							Low	Low
73	104	96-104	5.94			87-95	1.26	Mix	Low (Middle)
74	107	100-107	2.89			92-99	3.11	Low	Low (Middle)
75	123	113-123	2.03			90-100	4.88	Low	Low (Middle)
76	123	116-123	2.15			93-100	4.26	Low	Low (Middle)
77	200	183-200	8.67	133-150	13-94	83-100	6.63	Cumulonimbus	Cumulonimbus
78	188	171-188	9.25	133-150	17.00	83-100	9.50	Cumulonimbus	Cumulonimbus
79	172	143-172	37.94	113-142	15.34	71-100	1.95	Mix	Cumulonimbus
80	206	180-206	11.69	124-150	29.23	74-100	5.13	Cumulonimbus	Cumulonimbus
81	190	171-190	11.26	131-150	23.29	81-100	8.87	Cumulonimbus	Cumulonimbus
82	116	106-116	10.30			90-100	1.45	Mix	Mix
83	116	106-116	6.26			90-100	1.26	Mix	Mix
84	141	121-141	19.92			80-100	2.97	Mix	Mix
85	92							Low	Low
86	90							Low	Low
87	87							Low	Low
88	111	106-111	2.51			95-100	5.04	Low	Low (Middle)
89	99							Low	Low (Middle)
90	90							Low	Low
91	180	161-180	15.86	131-150	17.19	81-100	7.89	Mix	Cumulonimbus
92	109	98-109	16.20			86-97	10.20	Mix	Mix
93	96							Low	Low
94	181	169-181	19.83	138-150	32.18	88-100	9.14	Cumulonimbus	Cumulonimbus
95	190	174-190	12.04	134-150	25.26	84-100	12.86	Cumulonimbus	Cumulonimbus

Table 4 (cont'd)

SAMPLE NUMBER	MAXIMUM IR OBSERVATION	COLDEST CLUSTER OBSERVATIONS	COLDEST CLUSTER EDGE STRENGTH	MIDDLE RANGE OBSERVATIONS	MIDDLE RANGE EDGE STRENGTH	LOW RANGE OBSERVATIONS	LOW RANGE EDGE STRENGTH	DECISION PROCEDURE CLASSIFICATION	CLOUD-TRUTH CLASSIFICATION
96	116	103-116	6.13			87-100	2.45	Mix	Mix
97	101	92-101	2.64			82-91	2.72	Low	Low
98	92							Low	Low
99	91							Low	Low
100	93							Low	Low
101	88							Low	Low
102	90							Low	Low
103	96							Low	Low
104	91							Low	Low
105	189	171-189	12.94	132-150	22.10	82-100	10.56	Cumulonimbus	Cumulonimbus
106	180	165-180	24.88	135-150	33.73	85-100	14.07	Mix	Mix
107	84							Low	Low

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR - TR - 77 - 0640	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): TECHNIQUES FOR SEGMENTING INFRARED CLOUD COVER IMAGES.	5. TYPE OF REPORT & PERIOD COVERED: Interim
6. PERFORMING ORG. REPORT NUMBER: TR-515	
7. AUTHOR(s): JoAnn Parikh [redacted] Azriel Rosenfeld	8. CONTRACT OR GRANT NUMBER(s): F44620-72C-0062
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Computer Science Ctr. Univ. of Maryland College Park, MD 20742	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304/A2
11. CONTROLLING OFFICE NAME AND ADDRESS: Math.& Info. Sciences, AFOSR/NM Bolling AFB Wash., DC 20332	12. REPORT DATE: Mar [redacted]77
	13. NUMBER OF PAGES: 145
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Pattern recognition
Image processing
Scene analysis
Segmentation
Thresholding
Edge detection
Clustering

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report investigates three techniques for segmenting cloud cover images into regions of homogeneous cloud type. Two of these techniques select thresholds based on an analysis of the edge strengths of the borders of the above-threshold connected components (or of the coldest such component). The third technique selects thresholds based on cluster analysis of the infrared histogram, combined with a statistical feature analysis of the clusters in the image domain.